

© 2010

Mentor Cana

ALL RIGHTS RESERVED

OPEN ACCESS REPOSITORIES IN THE CULTURAL CONFIGURATION OF  
DISCIPLINES: APPLYING ACTOR-NETWORK THEORY TO KNOWLEDGE  
PRODUCTION BY ASTRONOMERS AND PHILOSOPHERS OF SCIENCE

by

MENTOR CANA

A Dissertation submitted to the  
Graduate School-New Brunswick  
Rutgers, The State University of New Jersey  
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in

Communication, Information and Library Studies

written under the direction of

Dr. Marija Dalbello

and approved by

---

---

---

---

New Brunswick, New Jersey

May, 2010

## **Abstract of the dissertation**

# **OPEN ACCESS REPOSITORIES IN THE CULTURAL CONFIGURATION OF DISCIPLINES: APPLYING ACTOR-NETWORK THEORY TO KNOWLEDGE PRODUCTION BY ASTRONOMERS AND PHILOSOPHERS OF SCIENCE**

By MENTOR CANA

Dissertation Director:

Dr. Marija Dalbello

This qualitative study provides an understanding of the role of self-archived disciplinary open access repositories in the cultural configuration of scholarly disciplines. It examines the implications of the technological and organizational layers of access tools and open access repositories and researchers' lived experiences and perceptions layer on researchers' localized knowledge production context and the construction of disciplinary knowledge production contexts. The actor-network theory, which posits that technological and social actors reciprocally affect each other, is applied to compare and contrast the information practices of two groups of researchers: the use of arXiv by astronomers, and the use of PhilSci by philosophers of science. Six astronomers and five philosophers of science were identified through purposeful selection. The interviews with the researchers were conducted over a period of five months, ranging in length between 40-75 minutes. Primary documentary evidence, describing open access repositories and access tools, is also used for the analysis. The findings show that the open access repositories, the access tools, and researchers' individual knowledge production contexts are co-constructed as researchers search, discover and access

scholarly artifacts. Open access has impacted researchers' knowledge production by realigning the existing processes and by instigating the emergence of new actors and constructs. Four themes emerge as researchers articulate their perceptions about the value and the role of open access: impact on scholarly process, impact on scholarly output, integration with scholarly context, and democratization of the scholarly discourse. Congruent with the domain-analytic approach, two distinct socio-technological models emerge. Astronomers perceive arXiv as important and critical in their scholarly information practices, with a central role in their discipline. However, Philosophers of science perceive PhilSci as having a limited value in their scholarly information practices and rather minimal role in their discipline. The properties of disciplinary cultures, such as the mutual dependence between researchers and the task uncertainty in a specific discipline, are implicated in the appropriation of the open access repositories and access tools at individual and disciplinary level. The socio-technological co-constructionist approach emerges as a viable theoretical and methodological framework to explicate complex socio-technological contexts.

Dedication

To my family

## Acknowledgement

The road to the finish line for my Ph.D. has been a challenging quest, always balancing the time and attention between work, family and continued reading and research.

I like to thank my family for their support while working on my papers and research.

I would also like to thank the School of Communication and Information (SC&I) faculty from whom I have learned a great deal while working on my Ph.D. It is here that I learned and developed the notion of the interconnectedness of the many actors that surround us in everyday life.

Especially, I like to thank my advisor Dr. Marija Dalbello for her thoughtful advice, intellectual and practical guidance, and for her patience. She empowered me to see deeper into the world that surrounds us, going beyond the superficial.

## Table of Contents

Abstract of the dissertation .....	ii
Chapter 1. Introduction.....	1
Chapter 2. Background and literature review .....	11
2.1 Scholarly Communication .....	11
2.2 Open Access .....	13
2.3 Knowledge Production.....	16
2.4 From paper to electronic publishing .....	17
2.5 Repositories, access tools, and researchers' information practices .....	19
2.5.1 Repositories, information environments and Digital Libraries .....	21
2.5.2 Access tools: Citation and abstract indexes, and scholarly search engines .....	23
2.5.3 Copyright Transfer Agreements (CTAs) .....	25
2.5.4 Researchers' information practices and access modes.....	26
2.5.5 Completing the picture .....	29
2.5.6 The OA in access tools .....	31
2.6 Individual and disciplinary knowledge production contexts and cultures .....	33
2.7 OA implications for the scholarly publishing process .....	36
2.8 Defining article accessibility, discoverability and visibility .....	39
2.8.1 Article accessibility .....	39
2.8.2 Article discoverability .....	40
2.8.3 Article visibility .....	41
2.8.4 The role of accessibility, discoverability and visibility.....	42
2.8.5 The role of CTAs for visibility, discoverability and accessibility .....	43
2.9 Critical actors as identified in the literature .....	45
2.10 Key terms.....	46
Chapter 3. The Research Problem.....	49
3.1 Research Questions .....	52
Chapter 4. Theoretical framework: Actor-Network Theory.....	54
4.1 The overall theoretical orientation – co-constructionist approach as the meta-theoretical direction for the study .....	55
4.2 Methodological implications of socio-technological analysis.....	56
4.3 Consideration of alternative theories and frameworks.....	57
Chapter 5. Methodological perspectives .....	61
5.1 Grounded theory and phenomenology .....	61

5.2	Semantic Elements of the Actor-Network Theory .....	62
5.3	Links and associations in building network topologies .....	69
Chapter 6.	Methods and methodological considerations.....	72
6.1	Research process congruency .....	72
6.2	Research design .....	75
6.2.1	Methodological approach .....	78
6.2.2	Data collection approach .....	79
6.2.3	Selection of participants .....	81
6.2.4	Analysis approach .....	82
6.3	The Research Process.....	86
6.3.1	Data collection .....	86
6.3.2	Identifying OA repositories .....	87
6.3.3	The research participants .....	94
6.3.4	The interview process.....	100
6.3.5	Documentary evidence: identification and analysis procedures .....	104
6.3.6	Software for qualitative analysis .....	109
6.4	Interviews: coding and analysis procedures .....	110
6.4.1	Memos and notes.....	110
6.4.2	Interview coding procedures.....	111
Chapter 7.	The emergence of four themes.....	114
7.1	Impact on scholarly process .....	116
7.2	Impact on scholarly output .....	117
7.3	Integration with scholarly context .....	118
7.4	Democratization of scholarly discourse .....	118
Chapter 8.	Analysis and interpretation of individual arXiv participants.....	120
8.1	Participant A1 .....	122
8.2	Participant A3 .....	135
8.3	Participant A17 .....	145
8.4	Participant A18 .....	155
8.5	Participant A19 .....	162
8.6	Participant A20 .....	172
Chapter 9.	Analysis and interpretation of individual PhilSci participants .....	180
9.1	Participant Ph10.....	180
9.2	Participant Ph24.....	185
9.3	Participant Ph26.....	194



9.4 Participant Ph29 .....	199
9.5 Participant Ph32 .....	204
Chapter 10. Comparing and contrasting the lived experiences and perceptions amongst and between the arXiv scientists and PhilSci scholars .....	214
10.1 arXiv scientists – group level experience .....	214
10.2 Differentiating and unique experiences amongst the arXiv scientists .....	223
10.3 PhilSci scholars – group level experience .....	226
10.4 Differentiating and unique experiences amongst the PhilSci scholars .....	231
10.5 Comparison of lived experiences and perceptions between astronomers and philosophers and science .....	233
Chapter 11. The role of the open access repositories and the access tools in researchers' knowledge production processes .....	241
11.1 The organizational perspective .....	242
11.1.1 arXiv (organizational) .....	243
11.1.2 ADS (organizational) .....	247
11.1.3 PhilSci (organizational) .....	250
11.1.4 Protocols and standards .....	252
11.2 The technological perspective .....	254
11.2.1 arXiv (technological) .....	255
11.2.2 ADS (technological) .....	258
11.2.3 PhilSci (technological) .....	261
11.2.4 EPrints .....	262
11.2.5 Google Scholar .....	263
11.2.6 Protocols and standards .....	264
11.3 The symbiotic relationships between the organizational and the technological layers .....	266
11.4 The impact of the technological and organizational layers on articles' visibility, discoverability and accessibility .....	275
Chapter 12. Study findings: contextualizing the individual and the disciplinary information practices .....	277
12.1 Answering the Research Questions .....	281
12.2 Socio-technological implication for the four themes .....	308
12.2.1 Impact on knowledge production process .....	309
12.2.2 Impact on knowledge output .....	310
12.2.3 Integration with the broader context .....	311
12.2.4 Democratization of the scholarly communication .....	312
12.3 Co-construction of the individual and disciplinary contexts .....	314

Chapter 13. The value of ANT for this study .....	317
Chapter 14. Evaluation criteria .....	322
Chapter 15. Conclusion .....	328
15.1 Implications and study importance .....	332
15.1.1 Contribution to LIS and scholarly communication .....	332
15.1.2 Contribution to theory and methodology.....	336
15.1.3 Contribution to practice .....	337
15.2 Study scope and limitations .....	339
15.3 Propositions .....	340
15.4 Further research .....	343
Appendix A: Types of research studies about OA .....	346
Appendix B: OA repositories selection process.....	349
Appendix C: Request for Participation (RFP) .....	352
Appendix D: Informed Consent Form.....	354
Appendix E: Pre-interview e-mail questionnaire .....	356
Appendix F: Interview questions .....	357
Appendix G: Interview Guide .....	358
Appendix H: The emergent categories and the codes representing the four main themes .....	362
Appendix I: Summary of perceived values and implications for each theme .....	372
Bibliography.....	381
Curriculum Vita.....	388

## List of tables

Table 1: Medium (print, electronic) vs. OA and NOA approaches .....	17
Table 2: Selected disciplinary repositories .....	89
Table 3: Brief statistics for the selected repositories .....	94
Table 4: arXiv and PhilSci participant characteristics.....	99
Table 5: Document analysis sources .....	107
Table 6: Summary of main categories, themes, and concepts and their sub-categories	115
Table 7: Summary of OA properties and their relation to the four themes as perceived by A1.....	134
Table 8: Summary of OA properties and their relation to the four themes as perceived by A3.....	144
Table 9: Summary of OA properties and their relation to the four themes as perceived by A17.....	154
Table 10: Summary of OA properties and their relation to the four themes as perceived by A18.....	161
Table 11: Summary of OA properties and their relation to the four themes as perceived by A19.....	171
Table 12: Summary of OA properties and their relation to the four themes as perceived by A20.....	179
Table 13: Summary of OA properties and their relation to the four themes as perceived by Ph10.....	184
Table 14: Summary of OA properties and their relation to the four themes as perceived by Ph24.....	193
Table 15: Summary of OA properties and their relation to the four themes as perceived by Ph26.....	199
Table 16: Summary of OA properties and their relation to the four themes as perceived by Ph29.....	203
Table 17: Summary of OA properties and their relation to the four themes as perceived by Ph32.....	212
Table 18: Synthesized list of properties for each theme as perceived by the arXiv scientists .....	215
Table 19: Synthesized list of properties for each theme as perceived by the PhilSci scholars.....	227
Table 20: Synthesized list of properties for each theme as perceived by the arXiv scientists and PhilSci scholars .....	234
Table 21: Model comparing lived experiences and perceptions .....	238
Table 22: Comparison of the technological layers of features and functional capabilities of arXiv and PhilSci.....	267
Table 23: Comparison of the organizational layers of intended roles and values of arXiv and PhilSci.....	269
Table 24: Comparison of the technological layers of features and functional capabilities of the access tools used by arXiv and PhilSci participants .....	271
Table 25: Comparison of the organizational layers of intended roles and values of the access tools used by arXiv and PhilSci participants .....	273
Table 26: High level summary and comparison of lived experiences and perceptions	284

Table B1: List of all 17 ROAR OA disciplinary (cross-institutional) repositories for the US .....	349
Table B2: List of OpenDOAR disciplinary repositories for the US .....	350
Table I1: Mapping and summary of the perceived values and implication of the properties for each theme for the arXiv group .....	372
Table I2: Mapping and summary of the perceived values and implication of the properties for each theme for the PhilSci scholars .....	377

## List of illustrations

Figure 1: Technology view of the process of author's interaction with OA repositories via access tools .....	20
Figure 2: Socio-technological view of the process of author's interaction with OA repositories and access tools.....	30
Figure 3: A conceptual framework of researchers' information access modes, access tools, and repositories .....	51
Figure 4: Interactions as triggers for scholar's lived experiences and perceptions .....	85
Figure 5: Performative model based on lived experiences and perceptions, technological and organizational layers.....	280
Figure 6: A model for lived experiences and perception layer based on the interviews with researchers .....	286
Figure 7: Relationship between disciplinary culture contexts and the technological and organization layers.....	301
Figure 8: Socio-technological model of the arXiv scientists .....	306
Figure 9: Socio-technological model of the PhilSci scholars .....	307

## **Chapter 1. Introduction**

The communication and publication facets of the Internet were embraced by the scholarly community mainly due to the inherent process efficiencies made possible by the electronic networked distribution capabilities enabled by the Internet. The socio-technological nature of scholarly communication as a structure of interconnected technological and social entities, and the unique and the novel ways in how various entities in the scholarly communication ecosystem may be interconnected and enacted to enhance the scholarly communication processes have enabled the emergence of new phenomena. One such phenomenon is the emergence and the construction of self-archived repositories where researchers can deposit their pre-prints and peer-reviewed articles for others to freely read them unhindered by the barriers put in place by commercial publishers. This newly developed context for the circulation of knowledge, as experienced and perceived by the researchers, includes at least the researchers (as authors and readers), a new set of tools that enable researchers to discover and access scholarly artifacts deposited in these repositories, and the actual repositories. Each of these elements that constitute the new context, inform and are informed by the local and global networks, thus play an important role in the emergence of new disciplinary frameworks for the production of knowledge that are co-constructed together with the new context.

As part of the knowledge production process, researchers as authors link to each other with the intent to build upon each other's work. Researchers seek out and reference other scholarly work (articles, books, reports, etc.) utilizing ideas, thoughts and theories upon which they further expand, apply them in different settings, or might challenge

them entirely (Littlejohn, 1992, p.21). The production and circulation of scholarly articles, viewed through the meta-theoretical framework of actor-network theory (ANT), is situated within the ecosystem which is constructed by actors (i.e., such as researchers as authors, access tools and article repositories), and a set of relationships that hold the ecosystem together. The relationships between the technological and the social actors form an ecosystem that can be explained and understood using the actor-network theory that posits that technological and social actors reciprocally inform each other to build and maintain highly dynamic but relatively stable and sustainable networks through institutional and disciplinary cultures.

For example, for the duration of the knowledge production process of a specific article, a researcher enacts a dynamic set of links connecting him with other researchers via the searching process for scholarly materials—these links are mediated by the repositories and the tools that link the researchers with the repositories. Thus, the produced article as a knowledge artifact takes a set of these dynamic links and makes them static in the form of citations, references, bibliography, indexes, etc.

As previously stated, in the context of scholarly communication, actors can be human or non-human such as computers, information systems, social structures, information artifacts, document repositories, digital libraries, etc. Some of these actors are immediately and locally linked to the scholar's production environment such as access tools, self-archived repositories, articles, and journals; others are part of the larger socio-cultural and technological context that is linked to the scholar's production environment in the form of information access behavior determined by such frameworks as disciplinary norms and cultures, Copyright Transfer Agreements (CTA), tenure, pre-

publication policies, funding, peer recognition, trust, technological frames (the software and applications that are used to build the access tools and repositories), etc. The difference between the local actors and the actors that constitute the larger socio-cultural and technological context is that the local actors are very visible and tangible to the researchers in their everyday scholarly work. The global actors are instead somewhat invisible from the perspective that they are not necessarily used in researchers' everyday work, but they do structure the local networks in significant ways even though their importance might not be obvious without deeper understanding how they are linked to the actors that constitute the local network.

Among these elements, there is empirical evidence for the importance of open access (OA) repositories in scholarly communication. Research shows that the number of available OA repositories of peer-reviewed articles has been increasing over the past few years (Brody, May 2006). Various citation impact studies have shown that peer-reviewed articles made available as OA have greater citation impact (Antelman, 2004; Harnad, May 2006). In addition, the number of peer-reviewed articles deposited in OA repositories has also increased over the past few years (Brody, May 2006). These findings indicate that OA repositories and the artifacts deposited therein are increasingly being used by researchers in their knowledge production and are becoming part of the scholarly publishing process. Thus, considering the increasing importance of open access () in scholarly communication, this study aims to understand the implication of open access repositories in researchers' knowledge production processes and more broadly their implication in the realignment of the disciplinary norms and cultures.



In parallel to the self-archiving of knowledge artifacts (especially of peer-reviewed articles) in OA repositories, hundreds of OA journals have been established so far to supplement the traditional proprietary publishing based journals with an open access alternative. The aim of the OA journals is to provide their content free of charge to anybody that wants to access them. Proprietary publishers have recognized the value OA journals provide to the scholarly community and have converted some journals into OA or into hybrid publications where articles become OA after some period of time. The dynamics of OA journals are different from the intent of this study and are not addressed as part of this study.

In exploring this framework for knowledge production, Chapter 2, Background, is organized as follows. *First*, based on the current literature on scholarly publishing, some critical challenges are identified resulting from the transformation of print publishing by networked scholarly publishing supported by the emergence of the Internet. *Second*, self-archived repositories of peer-reviewed articles and pre-prints and the open access movement are identified as relevant dimensions of scholarly publishing. *Third*, the knowledge production process that is the focus of this study is contextualized and scoped to encompass the processes enacted by the researchers in searching, discovering, and accessing articles to be used in the production of a scholarly artifact. *Fourth*, the shift from paper to electronic publishing and its implication for the OA phenomenon is explored both from technological deterministic and social deterministic perspectives. *Fifth*, a set of access tools are identified that serve as mediators between researchers and repositories. They include abstracts and indexes, references, specialized search engines such as Google Scholar and Scirus, as well as Copyright Transfer Agreements (CTA)

that provide the legal agreements between researchers and journal publishers and determine whether a pre-print and/or a post-print of an article can be published in self-archived open access repositories. Also, the role of researchers' information practices as manifested by the humanities scholars and scientists are identified as researchers' properties that could have critical implications for researchers' access to self-archived repositories. *Sixth*, researchers' individual and localized knowledge production contexts are positioned and related within the disciplinary knowledge production context. *Seventh*, a number of potential implications of OA for the scholarly publishing process are identified, starting with the core supposition that by the practical implementation of the OA principles (electronic distribution of content, free and unrestricted access, author retains rights), the peer reviewed scholarly work will be disseminated as widely as possible and be freely accessible by researchers and other interested individuals alike. *Eighth*, constructs that can be used to understand and describe the outcome of researchers' interaction with open access repositories are explored, described and summarized, with visibility, discoverability, and accessibility identified as potential constructs for the context of disciplinary self-archiving in open access repositories. *Ninth*, critical actors relevant to researchers' interaction with self-archived open access repositories are identified and summarized. *Tenth*, key terms are summarized and defined.

Chapter 3, The Research Problem, presents a conceptual framework of the study suggesting potential actors and their properties, followed by the research questions.

The description of the problem is followed by the theoretical and methodological frameworks in Chapter 4, where the actor-network theory (ANT) is introduced and

described with its methodological implications at three different levels. First, as a high level meta-theoretical thinking it suggests the co-constructionist approach because a set of actors, whether social or technological, are presented and expressed at a congruent interpretive level, and are associated and linked to each other. Thus, it empowers the investigator with a conceptual framework and terminology to develop a network topology of interconnected social and technological actors. Second, it is applied as a guide to knowing where to look and how to look in building a socio-technological network and especially identifying associations and links and tracing them. Third, it is applied as a mechanism to analyze the technological and the social together, instead of analyzing the social separately, as that leads to reductionism and comparison of conclusions with impoverished results. With the technological and social being analyzed together at the very detailed level, a more granular view of the co-constructionist dynamics will emerge.

Chapter 5, Methodological perspectives, presents the methodological perspective of the study. Here ANT is described in greater detail where foundational texts such as Akrich and Latour (1997), Bijker and Law (1997), Latour (1999), Latour (2005), Law (1999), Law and Bijker (1997), Law and Callon (1997), and Law and Hassard (1999), are introduced in relevance to the networked socio-technological co-constructionist approach.

While Kling, McLim and King (2003) have demonstrated the value of ANT specifically related to information science, other researchers have shown the value of ANT in wide variety of disciplines such as politics, arts, economics, power relations, science, organizational studies, medicine, technology, etc. Kling, McKim and King

developed Socio-Technical Interaction Networks (STINs) as a framework to analyze various forms of scholarly communication forums. One of the forums analyzed and modeled as a STIN is arXiv (p. 58-61), one of the earliest open access repositories of pre-print and post-print articles. Van House (2003) presents one of the closest studies that attempts to explicate the socio-technological dynamics of digital libraries (DLs) through the actor-network theory. However, apart from the identification of some actor-network vocabulary that can be helpful in understanding how various actors are related to DL contexts, Van House does not scrutinize those actors and relationships in great detail—perhaps because actor-network is not a central theme in her article. Abramson (1998) directly uses the translation capability provided by ANT in addressing the “national questions” of Canada and Quebec. ANT helps to delineate between micro and the aggregated macro (organizations, nations, universities, etc.) actors, stating that macro actors are only constructs that “speak” and are “spoken” upon, therefore able to perform and be performed upon. Bassi (1997) uses ANT to explain the relationships between intellectual capital and knowledge management, by emphasizing on the importance of training and the importance of performance professionals as actors that are able to inscribe and thus bring forth the power of intellectual capital. Miettinen (1999) contrasts and compares activity theory and ANT as means to understanding and explaining technical innovations situated in the nature and society dualism. It suggests that ANT is not the most appropriate approach for studying technical innovations. Wick (2001) argues that a successful implementation of knowledge management (KM) concepts is conditioned upon the ability to measure the outcome of KM. ANT's “inscription” and “translation” concept are apparent in the design of web-enabled monitoring tool called

Intellectual Capital Dashboard (ICD) and its associated metrics: contractor usage, customer and employee satisfaction, employee turnover, number of innovations, etc. Ducheneaut (2005) applies ANT to understand the dynamics as socialization, as learning and as a political process, between the developers of open source software (OSS) and the materials they deploy collectively around specific projects to produce freely available open source software.

In Chapter 6, Methods and methodological considerations, the design of this study is presented, where phenomenology and grounded theory approaches complement the actor-network perspective, acting as a guide about where to look and how to look for various socio-technological actors that might emerge as relevant in researchers' interactions with open access repositories. On the methods side, grounded theory emerges as congruent data collection, data analysis and interpretation tool to help with the understanding and description of researchers' lived experiences and the process that is being enacted therein, as the researchers interact with OA repositories via a number of access tools in the knowledge production process. A set of actor-network semantic elements complement the interpretation and description. Actor-network's "inscription" and "translation" constructs enable impartial description of the dynamic of socio-technological contexts and their actors and relationships.

Unlike the usual application of grounded theory as a methodological tool that informs theory building inductively from data (Strauss & Corbin, 1998, p. 12-14), the application of grounded theory for this study is applied as a methodological tool for data collection, analysis and interpretation, with the goal to describe the dynamics of researchers' interaction with open access repositories and the disciplinary knowledge

production contexts. Instead of theory building, the study puts forward propositions that can be the basis for future studies.

In Chapter 7, The emergence of four themes, the initial analysis based on the open and axial coding is presented, with four themes of discourse emerging as containers of researchers' lived experiences and perceptions. The researchers perceive open access as implicated at individual level (impact on scholarly processes and knowledge networks) as well as disciplinary level (as democratization of scholarly process and integration with scholarly context).

The nature of the relationship amongst the four themes is further analyzed in chapters 8 through 11. The analysis and interpretation of the individual participants, astronomers and philosophers of science, is presented in Chapter 8 and Chapter 9. In Chapter 10, the participants are then presented and compared as two distinct disciplinary groups of researchers, namely astronomers that use the arXiv open access repository and philosophers of science that use the PhilSci repository, focusing on their common patterns of perception but also on their differences and unique lived experience and perceptions. In Chapter 11, the organizational and technological properties of the open access repositories and the access tools are presented and interpreted. Throughout chapters 7, 8, 9, 10, and 11, the grounded theory is applied with its' open, axial and selective coding to discover the main categories and concepts emerging from the data. The actor-network theory (ANT) is used in conjunction with the grounded theory methods, to help interpret the nature and dynamics of the relationships between the emergent concepts and categories across the lived experiences and perceptions layer, the technological layer, and the organizational layer.

In Chapter 12, the findings of the study are summarized by answering the Research Questions. The co-construction of the individual and disciplinary information practices are contextualized in relation to the four themes.

Chapter 13, The value of ANT for this study, elaborates about how ANT has specifically informed and guided this study from the broader perspective as well as with respect to specific methods used in conjunction with the grounded theory approach.

Chapter 14, Evaluation criteria, is a process of self-assessment about the applicability of the grounded theory for the investigation and explication of the research problem of this study.

Chapter 15, Conclusion, reflects on the dynamics of the overall study and discusses the implications for scholarly context, theory, methodology, directions for future research, and limitations. Also, a number of propositions are advanced based on the findings.

## **Chapter 2. Background and literature review**

This section presents the overview of relevant literature and provides the underlying rationale for conducting this study. It investigates discourse related to scholarly communication and the implications from the introduction of open access concepts. It also identifies some gaps in the discourse and identifies few potential constructs that can be helpful in studying researchers' experiences as they embark on a quest to knowledge production. This section concludes with the identification of critical and relevant actors that have been mentioned in the literature. A number of terms and definitions that will be used in this study are adopted and defined at the conclusion of this section.

### *2.1 Scholarly Communication*

Kling and Covi (1995) noted that the linking property of scholarly publishing as the primary task in the scholarly communication process “should be viewed as one part of the scholarly communication systems that connect authors and readers” (p. 3). The relatively stable scholarly publishing process has recently been affected by the discourse centering on two major socio-cultural issues: a) the “serials crisis” linked to an ever increasing subscription cost for libraries to acquire journals published by commercial publishers, and b) the “permission crisis” linked to the inability of the readers to access the bewildering amount of scholarly output due to copyright concerns that restrict what can be read and circulated (Kling & Callahan, 2003, p. 127; Okerson, 2000, p. 672; Van de Sompel, Payette, Ericksson, Lagoze, & Warner, 2004, Introduction section, para. 1). Further, a number of researchers have identified process inefficiencies with the



established scholarly communication system. The issue of latency between scholarly results and their publication (Van de Sompel et al., Introduction section, para. 1; Okerson, p. 672), both in print and electronic publishing, encapsulates the challenges with the process inefficiencies that in turn slows down the publication cycle and thus inhibits innovation. Schauder (1994) identifies the following comprehensive set of problems that have manifested themselves within the commercial journals system in the past two decades: a) the refereeing system suppresses new ideas, tends to favor article sources from prestigious institutions, causes undue delay in publishing articles, b) there are too many journals, c) journal subscriptions are too expensive, d) journals are not selective enough – poor quality articles get published, e) good articles remain unpublished because journals cannot cope, and f) research is old news by the time an article appears in the journal (p. 75-76).

The remedies for the challenges identified by Schauder (1994) have emerged as: a) the advent and low cost of the personal computer, b) the advent of information technology tools in the form of word processors and desktop publishing tools, and c) the advent of the networked communication, having manifested itself as the Internet, an ubiquitous communication and publishing technology. These technological changes, driven by socio-technological co-construction, have enabled researchers to use their time more efficiently by moving a portion of the knowledge production tasks away from print and paper by using electronic publishing capabilities. Before the Internet, these changes were isolated to the production task of preparing the manuscript. The same shift has induced changes with the journal publishers by streamlining their journal production into the more efficient electronic preparation, thus further increasing their profits, even

though the final product was still mostly circulated in paper format. Next, as the Internet became available and it became mainstream, publishers started to experiment with electronic publishing initially in the form of p-e-journals (“p-e” indicates primarily distributed in paper form but are also distributed electronically), and later with pure e-journals (originally distributed only in digital form), e-p-journals (primarily distributed electronically but may have very limited distribution in paper form), and p+e-journals (initiated with parallel paper and electronic editions) (Kling and Callahan, 2003, p. 134). The Internet has also enabled researchers to communicate amongst themselves without the mediation of the publishers, something that was not feasible prior to the Internet—possible only in a very cumbersome and limited way via postal mail, faxes, and scholarly conferences. The Internet has brought new dimensions to scholarly communication via discussion lists, electronic collaboration communities, enhanced invisible colleges, self-archive repositories, etc. As a result, the researchers have steadily realized that even though commercial publishers add a tremendous value to the scholarly publishing process, they can be augmented or even replaced by alternative publishing models. It is important to note that the most critical tasks in scholarly publishing, such as article writing, the peer-review, and editing processes, are mostly performed by the scholarly community itself. Therefore, it is important to study the internally driven, endogenic processes represented by the OA and self-archiving communities.

## 2.2 *Open Access*

The historical development of scholarly publishing has positioned commercial publishers as mediators to funding, tenure, hiring, grants, and recognition, vested upon them by their role to mediate between authors and readers—who may be the authors

themselves. Many journal publishers are commercial entities whose commitment is to the bottom line (making profit) and secondary to disseminating the articles quickly and as widely as possible. The discourse around the rising costs and inflexibilities of scholarly products to be used freely due to copyright restrictions enacted by the commercial publishers, has positioned the open access (OA) movement as a serious challenge to the well-established scholarly publishing practices. The conjunction of the broader OA movement, which claims that the free and unrestricted use of information will be a positive force for innovation and human progress, and the availability of enabling technologies at individual as well as organizational level that can address process inefficiencies, has enabled the emergence of the OA phenomenon in the scholarly publishing process. The most comprehensive OA initiative, the Budapest Open Access Initiative (BOAI, 2002) defines OA as "... the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds." Based on the BOAI definition, the aim of OA is to address both the serials crisis (by providing a cheaper alternative to commercial publishers through OA journals yet maintaining the quality provided by the peer review process), and to address the permission crisis (by enabling the widest possible distribution of scholarly articles via disciplinary self-archiving of peer reviewed articles that have already been published in commercial publications). In addressing the serials and the permissions crises, OA takes full advantage of the electronic means of knowledge production in the form of electronic networked distribution of scholarly articles in order to improve the process inefficiencies.

The arguments of OA supporters against the commercial publishers started as a dissatisfaction with their behavior of rising costs and stifling innovation due to very restrictive copyright rules—barriers to access and dissemination. But the appropriation of OA has also initiated a very practical set of processes (enablers of access) stemming from researchers' desire to have their knowledge disseminated as widely as possible, as well as the ability for them to read each others' work without inhibition:

Though the open access movement was galvanized by the price-gouging of commercial scientific publishing, it's not really about the money. It's an appeal to authors to make their scholarship freely available; and then an appeal to publishers not to interfere in this process, and eventually to put their publications online also. Already the open access movement in scientific, technological and medical publishing has lead to a number of positive outcomes: reduction in costs to libraries for these journals, a greater commitment to free access to readers, and even the prospect of realistic alternatives to the business model of commercial scholarly publishing. As a movement it is extremely robust, and gaining in force every day (Hunter, 2005, p. 13)

As a result of OA, the scholarly community can reap practical and immediate benefits. It is suggested that the more the work is circulated and made more visible and accessible, the more it will be read. This allows for better use of resources in the production of new knowledge, resulting in advanced pace of research and increased

innovation, with the basic assumption that individuals have Internet access from home, school or workplace.

### 2.3 *Knowledge Production*

Scholarly publishing, as an element of scholarly communication, encompasses the knowledge production process, knowledge circulation and knowledge re-circulation in the form of articles and other knowledge artifacts. As it relates to the production of scholarly peer-reviewed articles and other knowledge artifacts, knowledge production is the activity that is undertaken by the scholars and scientists as authors. There are two distinct parts to the knowledge production: a) the micro level or the actual construction of discourse and the writing of the article with the intricacies of organizing various disparate information and knowledge elements into a meaningful whole in the form of a knowledge artifact, and b) the macro process of knowledge production that feeds the micro level, that is the process of searching, discovering and accessing information to be used in the production of knowledge artifacts. This latter part involves the use of various types of repositories and access tools. Some examples of knowledge artifacts are articles, pre-prints, project reports, funding proposals, scholarly presentations, etc.

This study relates to the latter aspect of the knowledge production process where the OA phenomenon, by enabling articles to be accessed freely in OA repositories, has changed the dynamics of information seeking behaviors, the dynamics of knowledge acquisition processes, expectations and meanings, assumptions regarding the structure of knowledge networks, as well as the dynamics of information access behaviors and information practices instantiated by the researchers.

## 2.4 From paper to electronic publishing

Implementing a self-archived OA repository within the ecosystem of print and proprietary publishing meets a number of barriers that are hard to overcome due to the cost and inefficiencies in duplication and distribution of paper based articles. Table 1 shows a simple matrix of OA/NOA modes of circulation as they relate to print/electronic aspects of scholarly publishing. It shows that OA is feasible in relation to the knowledge production process where knowledge artifacts are deposited in digital format and distributed via electronic channels. Table 1 is not a comprehensive table, rather it is used as a guide that conceptually delineates the cross-sections of OA/NOA modes of circulation and the print/electronic aspects of scholarly publishing. Its purpose is not to describe nor categorize hybrid publishing environments that contain both OA and non-OA access knowledge artifacts.

Table 1: Medium (print, electronic) vs. OA and NOA approaches

	<b>Print publishing ecosystem</b>	<b>Electronic publishing ecosystem</b>	
		<b>DL</b>	<b>Non DL</b>
<b>OA approach</b>	<ul style="list-style-type: none"> <li>• OA repositories not feasible due to cost and inefficiencies in duplication and distribution of paper articles</li> <li>• Some forms of print pre-print culture, but without a central repository</li> </ul>	<ul style="list-style-type: none"> <li>• OA repositories feasible</li> </ul>	<ul style="list-style-type: none"> <li>• OA repositories feasible</li> <li>• Challenging due to lack of organizing framework provided by DLs</li> </ul>
<b>Non OA approach</b>	<ul style="list-style-type: none"> <li>• The traditional model of print culture</li> </ul>	<ul style="list-style-type: none"> <li>• Commercial model of publishing</li> </ul>	<ul style="list-style-type: none"> <li>• Closed commercial database systems, without a meaningful organized structure of articles</li> </ul>

Note: Digital Libraries (DLs) are socio-technological systems that enable the collection, organization, archiving and maintained of digital content in systematic, meaningful and efficient way.

The shift from print to the electronic publishing medium has facilitated the emergence of OA. The role of the publishing medium as one of the relevant actors in the

dissemination of articles has been explored both from technological deterministic and social deterministic perspectives. Representing the technological deterministic perspective, exploration of Meyrowitz's (1994) various medium theories suggests that the shift of the scholarly publishing process from print to electronic formats has induced changes into the process that made it feasible for OA repositories to emerge. Meyrowitz contributes these changes to the electronic medium's intrinsic and extrinsic properties and its distinct interplay with its environment. In comparison with electronic production of an article and its subsequent low marginal cost for creating additional electronic copies and the cost for electronic circulation and recirculation, paper copies of articles (as indicated in Table 1) are more costly to produce in great numbers and are more costly to distribute them widely. Next, organizing and managing paper copies of articles to establish a repository of thousands of articles faces the limitations inherent in the manual manipulation of paper-based indexes, references and other paper based access tools. If articles are in electronic form, a multitude of access tools and organizing structures can be used for the management and organization of the articles. Thus, an organizing framework is more feasible if articles are circulated in digital format. If the aim of a repository is to enable researchers as authors to access the articles they need for their knowledge production, it is clear that paper-based repository will be beneficial only to local researchers. However, these researchers might not see a value in a paper-based repository since they probably can access the articles they need at their local research library.

In contrast to the technological deterministic view of medium theories, Kling and McKim (2000) have used the social shaping of technologies (SST) framework arguing

that the shift from print to electronic publishing is induced by social structures. Kling and Covi (1995) further identify economics as a reason for moving from print to electronic modes of production and publication due to the reduced cost, speed of dissemination, as well as the ability to include other scholarly output with the articles.

These two perspectives, technological determinism and social determinism, when taken together are very complementary as both perspectives address specific aspect in the dynamics of the scholarly publishing process with the introduction of electronic tools for the creation, dissemination and management of scholarly works. Technological determinism enables us to understand and describes the social constructs that emerge as a result of technological innovations, such as the emergence of organizational structures to deal with availability and use of digital libraries for specific purpose. Social determinism enables us to understand the technological constructs that emerge as a result of socially induced changes in specific context. For example, email as a technology that enables one-to-one communication, has been appropriated to also enable one-to-many communication. Groups of people have shared ideas, opinions and have discussed as a group in a town-hall style setting for thousands of years. Listservs or mailing lists software is the technology that enables this type of many-to-many communication. It has been developed using e-mail technological capabilities as its basis, driven by the need to establish town-hall style conversations.

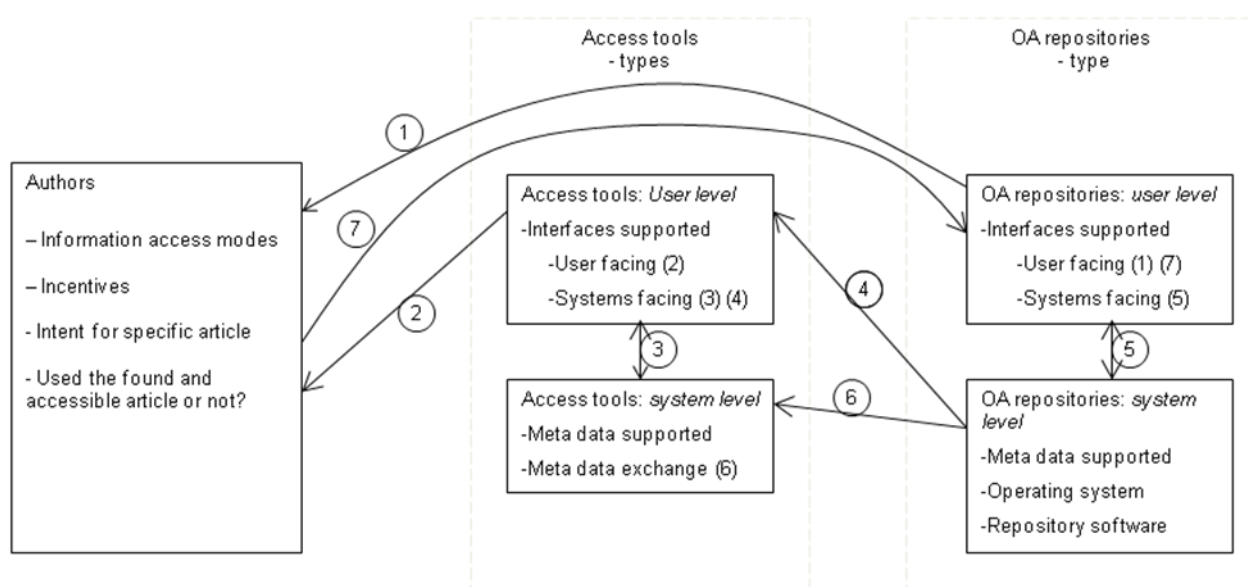
## *2.5 Repositories, access tools, and researchers' information practices*

Scholarly articles that are deposited in meaningfully organized collections with the intent to be easily discovered and easily accessed, greatly enhance researchers' tasks in knowledge production. Unlike the challenges faced with paper based publishing,



electronic publishing of articles makes it feasible to build and maintain repositories of articles. A number of technological capabilities are deployed to enable these repositories. The repositories are software systems built with the capability to store, organize, maintain and make available the digital articles through a system or user interfaces.

Figure 1: Technology view of the process of author's interaction with OA repositories via access tools



Note: 1 - 7 represent interfaces between the authors and the systems, and between the systems.

The technological capabilities of the enabling systems are split between the system level capabilities in the form of software (EPrints, DSpace, Greenstone, Open Journal Systems (OJS), etc.) deployed by an organization or institutions that establish an information environment, and the user level capabilities that are deployed and used by the researchers in the process of knowledge production, shown in Figure 1 with interface (1) and (2). Examples of user level capabilities are Google Scholar, Google Desktop, DL-Harvest (open access aggregator and a search engine for the Library and Information Science discipline), OAIster, etc. These two sets of technologies, the system level and

the user level, need to communicate with each other seamlessly, as shown in Figure 1 with interfaces (3), (4), (5) and (6), in order to benefit the scholar in the quest of knowledge production, and in knowledge acquisition from information environments such as electronic journals and digital libraries (DLs) at journal level, and pre-prints and post-prints at article level. In contrast to these access enabler tools at the technology level, it is critical to note that the legal framework in the form of Copyright Transfer Agreement (CTA) acts as a gate at article level determining whether an article can be deposited in a repository or not. Thus, the CTAs are inhibitors to OA—they act as barriers between readers and peer-reviewed articles that have already been published by a commercial publisher.

#### 2.5.1 Repositories, information environments and Digital Libraries

The information processing requirements for the systematic and predictable access to scholarly artifacts is that they are collected, organized, archived and maintained systematically in a meaningful and efficient way. Digital Libraries (DLs) are the technology that captures, disseminates, makes accessible and preserves over time the varying nature of scholarly output; not only articles, but also tables, raw data, pictures, algorithms, actual code, interfaces, and digital artifacts (Van de Sompel et al., 2004). In this sense, pre-print and post-print archives, portals to collections of scholarly work, disciplinary and institutional repositories, electronic journals, and even personal Home Pages of researchers with some organizational overlay to the content, can be considered DLs. The content is usually a research article, but also can contain tables, raw data, pictures, figures, references, databases, metadata, and other indexes and abstracts that may or may not be linked to any specific research article.

In considering the role of various information environments that manifest themselves in various DL instantiations, both socio-cultural and technological aspects of digital libraries need to be considered. OA repositories of self-archived articles are DLs that link researchers amongst themselves (as producers of knowledge and as readers), as well as linking researchers to the institutional or disciplinary context where the repositories are embedded. Repositories also link researchers with the technologies that these systems are built with, although these links are not necessarily visible to the researchers and might not appear to be critical to the knowledge production process. Bijker and Law (1997) have shown via many examples that technologies are not produced in isolation, rather, “our technologies mirror our societies. They reproduce and embody the complex interplay of professional, technological, economic, and political factors” (p. 3). Similarly, Bijker and Law further elaborate and describe how technologies affect the social structures within which they are embedded: “For the social is not exclusively sociological. In the context of technology and its social shaping, it is also political, economic, psychological—and indeed historical” (p. 4).

A DL system is partially a software technology, and like any other

technology does not spring, *ab initio*, from some disinterested fount of innovation. Rather, it is born of the social, the economic, and the technical relations that are already in place. A product of the existing structure of opportunities and constraints, it extends, shapes, reworks, or reproduces that structure in ways that are more or less unpredictable. And, in so doing, it distributes, or redistributes,

opportunities and constraints equally or unequally, fairly or unfairly  
(Bijker & Law, 1997, p. 11)

This approach is also congruent with Borgman's (1999) description of DLs as both social and technological entities: "digital libraries are a set of electronic resources and associated technical capabilities for creating, searching and using information ... constructed, collected and organized, by (and for) a community of users, and their functional capabilities support the information needs and uses of that community" (p. 234).

#### 2.5.2 Access tools: Citation and abstract indexes, and scholarly search engines

In addition to scholarly publishing studies that address the serials and permission crises, Palmer (2005) emphasizes the need to study other aspects of scholarly communication by considering the access resources that enable a scholar to find the raw materials for the knowledge production process. Palmer defines access resources as "... the means by which researchers identify recognized scholarly works within and, sometimes more importantly, outside their field of research" (Palmer, 2005, p. 1141). Palmer also notes the limitation of paper oriented access resources that are restricted to bibliographies, references, and other description of collected information (p. 1141). This limitation due to the disconnected nature of the paper based access tools that require manual handling on a case-by-case basis is also evident in the electronic supported knowledge production environment that has not necessarily taken full advantage of the new tools. It has rather transplanted the old with the new without taking the advantage of the electronically networked environment that can meaningfully interlink various access

tools and further electronically link them to the actual articles and other knowledge artifacts.

It is important to recognize the value of Palmer's (2005) observation that researchers' information seeking behavior during the knowledge production process is mediated by various access resource tools and as such should provide some guidance about how to organize the knowledge output: "Scholars' modes of access and their 'working' and 'implicit' assemblages of information represent what researchers actually do when gathering and working with research materials and therefore provide a useful framework for the collection and organization of access resources in research libraries" (p. 1140). Palmer defines working assemblages as those materials that emerge from the informal connections between researchers who communicated to each other in the process of gathering information for their work (p. 1142). A thematic collection of materials collected and accumulated during the study of a specific theme is an example of a working assemblage (p. 1148). Implicit assemblages are defined as the access resources that emerge from the intellectual connections between authors that can be embedded and can be traced through citations among bodies of literature (p. 1142). There is an explicit suggestion that access resources and access tools are enablers of access to scholarly work and that not all organization of materials is equally supportive. Rather, the information environments where scholarly work is deposited, such as self-archived repositories, need to be designed with "discoverable" capabilities and functions compatible with researchers' information practices in terms of resources and services, enabled by the shift towards digitization of the access resources (p. 1140). Palmer further suggests that access resources created by the researchers such as references,

tables and figures of work in progress, be also treated as publishable items that can be deposited in scholarly repositories for access by other researchers and that they should be maintained and organized for ease of access (p. 1150).

### 2.5.3 Copyright Transfer Agreements (CTAs)

The drive to make scholarly articles more accessible faces a critical barrier: most commercial publishers have required and still require that researchers transfer the copyright to the publishers by signing some form of copyright transfer agreement (CTA). In their functional content analysis study of publisher's copyright transfer agreements (CTA) of 80 publishers (7135 journals) in the context of copyrights and OA pre-prints and post-prints, Beier and Tschida (2003) found that "90% of the examined agreements ask for copyright transfer"(p. 3). Coleman and Roback (2005) have found that most Library and Information Science (LIS) journals do not have clearly defined and visible copyright transfer agreement and those that do have, have heavily restricted self-archiving (The Practice of Self-Archiving in LIS, para. 4). Thus, such articles remain "visible" only to those that are able to pay (individually or through institutional access) to read the article. While CTAs appear as inhibitors to wider article circulation because they limit the potential for articles to be seen by a wider audience, the move from print to electronic publishing is seen as an enabler for wider article circulation because it has enabled the creation of effective and efficient access tools (user-to-systems and system-to-system) for the discovery of scholarly articles and other digital objects, such as Google Scholar, Scirus, DL-Harvest, CiteSeer, etc. Challenged with the emergence of self-archived OA repositories, some commercial publishers have relaxed their CTA policies and will grant exemption to researchers who explicitly request to be

able to publish a pre-print and/or post-print version of their article on their Home Page, or institutional or disciplinary repository.

#### 2.5.4 Researchers' information practices and access modes

Palmer (2005) identifies two modes of researchers' information access behaviors depending whether researchers belong to humanities or sciences disciplines:

The work of humanities scholars is centrifugal; they work their way out into the information universe to collect a base of information in which deep inquiry can be performed through reading and writing. The information gathering process is relatively open-ended, moving outward from lead to lead (p. 1146)

The work of the scientists is centripetal with information being pulled back to the locally generated data and results to solve the problems and questions encountered in each stage of an experiment or project. Searching, collecting, and consultation are more targeted and endpoints tend to be more defined (p. 1146)

If we assume that various disciplinary repositories are built by and for the particular scholarly community itself, the information access modes identified by Palmer (2005) will also have implications for how the repositories are built and the level of interoperability and compatibility with the access tools. Humanities scholars would thus prefer repositories that are organized thematically and the access tools are utilized to find the articles in these repositories. These access tools and repositories are not necessarily linked to each other beyond superficial linkage—usually the association of content is carried mostly by the scholar to create the new article. For scientists, in

addition to enabling them to discover and access articles, the access tools and repositories play an additional role in the knowledge production. Due to the factual nature of sciences, access tools and repositories can be linked and associated at content level. For example, raw data about a particular phenomenon in astronomy can be linked directly to uncover new potential associations. Thus, some of the capabilities of access tools and repositories used by scientists are the ability to deposit and access entities within the articles, such as raw data, formulas, tables, etc. (Palmer, p. 1148).

Palmer and Cragin (2008) further contextualize researchers' information practices more broadly within the relevant disciplinary cultures. This article provides a survey of research related to how researchers approach and enact the knowledge production process from searching, discovering and accessing articles for their knowledge artifact, and up to the writing process—noting differences between disciplines. The article shows that researchers' information practices are interrelated with the disciplinary practices, and suggests that the role of information technology artifacts and tools for scholarly work can be examined and understood by understanding how researchers themselves develop domain or discipline specific information technology.

In a study about how humanities scholars and scientists approach the searching of electronic resources and their relation with the literature in their specific discipline, Bates (1994) shows that humanities scholars and scientists have distinct approaches to using electronic resources for their knowledge production. Bates notes that for scientists an electronic database is perceived more or less as a tool for finding resources for scholarship with the main data for research being produced in a laboratory environment or via experiments, whereas for humanities scholars the body of literature found in



libraries, and perhaps in electronic resources, is actually the data upon which humanities scholars ponder and use as they produce their knowledge artifacts. The notable difference pointed out by Bates is that libraries and the electronic systems can be perceived as scholars' laboratory environments: "The statement that 'the library is the laboratory of the scholar' holds much truth to it. For the scientist, the laboratory or the field is where the action of research takes place, while the library is where the results are archived. For the humanities scholar, those two functions of research and archiving largely take place in the same place, wherever the information is kept" (p. 6).

Similarly, in her study about how communication channels and digital resources are utilized by the members of a specific discipline in relation to the disciplinary social and knowledge production cultures, Fry (2006) builds on Whitley's (2000) theory that attempts to categorize scholarly disciplines and intellectual fields with respect to the mutual dependence between the researchers and the research task and problem uncertainty. The strength of Whitley's framework is that it relates the social aspects of an intellectual field (encapsulated as a mutual dependence between the researchers—defined by its functional and strategic dependence) and the epistemic aspect of an intellectual field (encapsulated as a task or research problem uncertainty—defined by its technical task and strategic task uncertainty). Fry expands Whitley's framework by introducing the digital practices of researchers in their knowledge production, with the intent to explicate the implications of the use and structuring of digital resources and digital tools on researchers' knowledge production processes (p. 303). Fry (2006) finds that disciplinary communities that exhibit high degree of mutual dependence in conjunction with low degree of research problem uncertainty (such as astronomy and

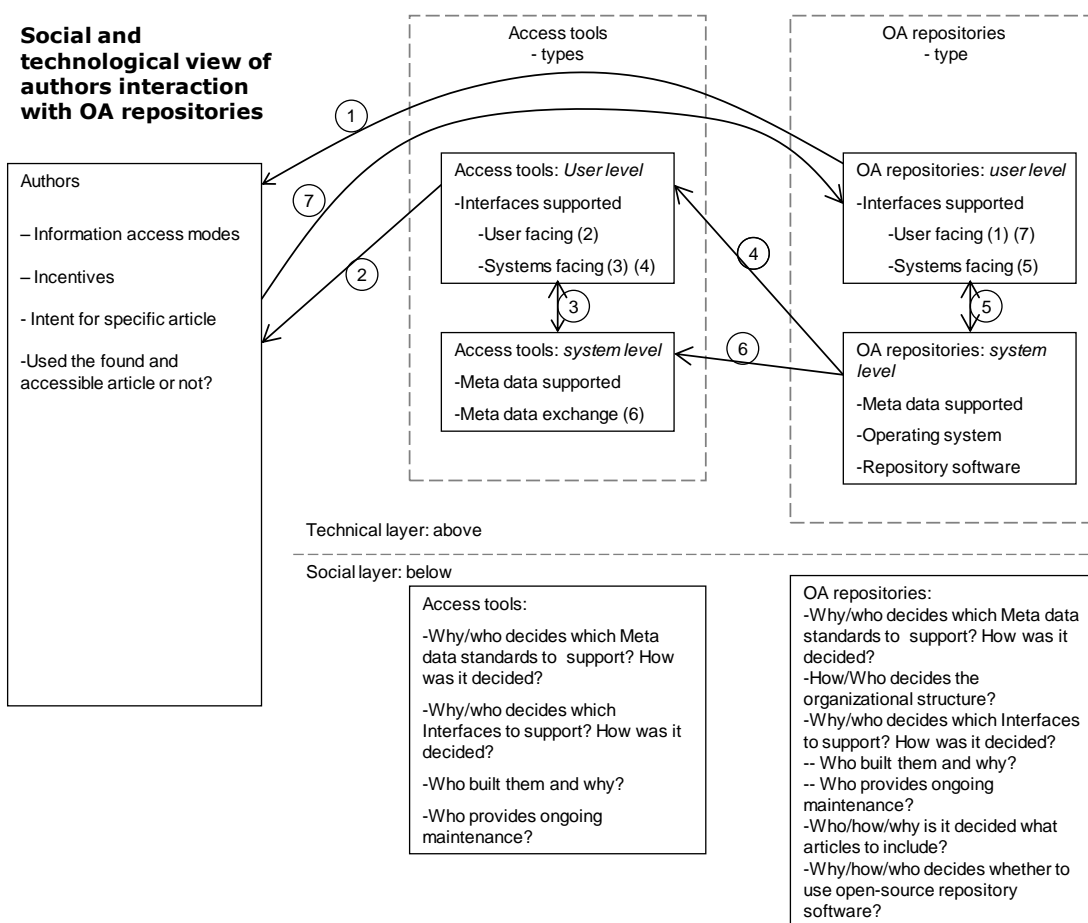
astrophysics), are more likely to utilize and co-construct digital tools and resources that will make their information practices more productive and more efficient. And, disciplinary communities that exhibit low degree of mutual dependence in conjunction with high degree of research problem uncertainty (such as philosophy of science), are less likely to utilize and co-construct digital tools and resources to the fullest potential (p. 299).

Thus, it is expected that these noted differences in information practices between humanities scholars and scientists will have implications on the dynamics that emerge as they interact with open access repositories.

#### 2.5.5 Completing the picture

Considering that repositories and access tools as technological systems are build and developed in social contexts (Bijker & Law, 1997, p. 11), for rich and deeper understanding of researchers' interaction with access tools and with OA repositories, Figure 1 can to be modified to account for the social context. This is accomplished in Figure 2 that shows a more comprehensive view of researchers' interaction with repositories—it accounts for the socio-technological co-construction of access tools and repositories. It is interesting to note that as researchers are using the access tools and repositories, they are probably not aware of the numerous socio-technological actors and processes that eventually factor into their lived experiences and perceptions.

Figure 2: Socio-technological view of the process of author's interaction with OA repositories and access tools



The user level experiences, such as browsing or searching the OA repositories either directly through their portals or indirectly through the use of access tools, is constructed through the technological capabilities of the access tools and the OA repositories. The technological capabilities by themselves are a result of the technological possibilities and the organizational and social structures that are enacted to manifest them in the form used by the researchers. For example, the organizing and management structure of a certain OA repository may not support all metadata standards and may not necessarily support all file types. This in turn will determine what is visible, discoverable and

accessible through the specific access tools and what is depositable in the OA repositories, thus affecting user experiences. The users of the systems may also impact the functionality of the access tools and the open access repositories through their feedback, and they may also use some features and capabilities in ways not intended by the designers and developers.

#### 2.5.6 The OA in access tools

The current state of the open access movement, as it relates to scholarly communication, takes a social perspective into account, with specific attention to the changing nature of the computing paradigm towards openness and interoperability. This paradigm operates with the assumption that the enabling access tools used by a scholar need to be interoperable and compatible with the software systems used to build the repositories. Interoperable access tools and repositories will ensure that the technology at user level (user-to-system communication) and system level (system-to-system communication) is not a barrier to access; rather it enables various access tools to communicate with the repositories in a meaningful way to exchange information and metadata in a format understood by both ends. In order to increase the interoperability and compatibility at user level and system level, the scholarly community has developed a set of open protocols and open metadata standards to enhance the access to the knowledge repositories through system-to-system services, as well as user-to-system services. The system-to-system capabilities enable the creation of value added services such as specialized crawling, harvesting, indexing, and abstracting. The user-to-system capabilities enable the creation of user facing services for more efficient, effective, and contextual discovery of scholarly works, thus helping in the process of wider circulation

and re-circulation. The creation of these tools is based on the open source software development model enabling the researchers to directly consult with the software developers in the design of the necessary tools.

The congruence and convergence between open source and open access and the concept of openness in general has been discussed in detail by Willinsky (2005), who is the prime developer of the Open Journal System - OJS (Willinsky, 2006, pp. 73-5). Building DLs with open source software only follows a logical process based on economics—the cost of commercially produced DL system will be a barrier to enabling open access journals. Considering that the rising cost of commercially maintained scholarly journals is a very critical factor for the evolution of the open access movement, the low cost of building a DL is a practical enabler. In addition to the economic feasibility provided by the open source software, the open standards—on which open source software bases its exchange, format and communication protocols—increase the interoperability by removing the complexity of interoperability between proprietary implementations.

At the forefront of many open access initiatives is the Open Archives Initiative (OAI) that “develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The Open Archives Initiative has its roots in an effort to enhance access to e-print archives as means of increasing the availability of scholarly communication” (OAI F.A.Q, What is the mission of the Open Archives Initiative?). A full list of tools based on the OAI protocols is available at <http://www.openarchives.org/tools/tools.html>. The OAI provides the OAI-PMH (Protocol for Metadata Harvesting), a protocol that defines and standardizes metadata

specific to scholarly publishing for self-archiving (including institutional and disciplinary repositories), for interoperability and for federated exchange of metadata across various open access archives and repositories. For example, in the Library and Information Science (LIS) field, DL-Harvest “a subject-based, open access aggregator and federated search service for LIS” (Coleman & Roback, 2005, Aggregating Open Access LIS, para.1) was developed in compliance with OAI-PMH. The most powerful aspect of the OAI-PMH is the ability for various post-print systems to exchange data with each other (as data providers), and enable value added services to be built outside of these systems, such as Google Scholar (user to system) and DL-Harvest (user to system, system to system) (Suleman & Fox, 2002). The value for the researchers results from the ability for the metadata and content from various repositories to be linked together so they can be searched and accessed in a more meaningful way. The effectiveness of Google Scholar and other similar tools to identify relevant resources based on a search query is enhanced by the availability of structured metadata about knowledge artifacts. Further, a tool like DL-Harvest that enables simultaneous searching of a number of repositories from one interface will be more difficult and not as effective without the ability to exchange metadata between various software systems that are used to build the repositories.

## 2.6 *Individual and disciplinary knowledge production contexts and cultures*

As has been shown in the previous sections, the information practices of researchers are embedded and intertwined with larger context. The scholarly production process (research, search and writing) and the information access tools used by researchers are discipline, domain or field specific (Palmer & Cragin, 2008). Thus, the structure or the

knowledge production context that is enacted as researchers produce knowledge comprises the scholarly information practices, researchers' localized and individual knowledge networks (the set of articles, journals, digital resources, libraries, etc. that are used for the production of a specific paper) and the broader disciplinary cultural norms and dynamics of the discipline. The amalgam of practice and process, individual knowledge networks, disciplinary knowledge networks, disciplinary cultural norms, and the empirical implications within a specific discipline is analyzed in great depth by Knorr Cetina (1999). She labels this amalgam as epistemic cultures and defines it as "the cultures of knowledge settings, and these appear to be a structural feature of knowledge societies" (p. 8), and further clarifies that studying epistemic cultures is not about the study of the structure and the construction of knowledge, rather it is about the study of construction of the "mechanisms" and the "machinery" of knowledge construction (p. 3). The use of the terms "mechanisms" and "machinery" by Knorr Cetina does not have deterministic connotation, rather, the term "mechanism" is used to denote the contextually constructed processes and the term "machinery" is used to denote the empirical, technological and social context that is co-constructed and enacted that enables the construction of knowledge. Knorr Cetina takes into account the empirical, technological and social "machineries" (p.24) of a discipline as important facets in trying to understand and describe the common patterns and differences between two disciplinary epistemic cultures. The empirical processes and structures used and applied by the researcher are out of the scope of this study.

This view and perspective is congruent with the intent of this study: the socio-technological and cultural embeddiness of researchers as they produce knowledge

artifacts manifests elements of process and localized knowledge networks at individual level, as well as process and global knowledge networks more broadly at disciplinary level. Disciplinary cultures operate by a set of norms and behaviors within which the individual researchers enact the knowledge production process. Thus, for this study, an individual knowledge production context is defined to denote the local and immediate processes and localized knowledge networks used by an individual researcher, and disciplinary knowledge production context is defined to denote disciplinary norms and cultures, which includes accepted practiced and accepted disciplinary knowledge networks in the specific discipline.

A constitutive element of a disciplinary knowledge production context is the disciplinary knowledge network, comprised of the knowledge artifacts available to the participants of the scholarly discipline through which and by which the epistemic elements of a discipline are defined. Researchers of the discipline use these artifacts in their knowledge production process and contribute to the same. At individual level however, researchers use a subset of the artifacts from the disciplinary knowledge network in the production of a specific scholarly work. This localized subset of knowledge artifacts for use by an individual researcher is the individual knowledge network. The form and shape of an individual knowledge network amongst other factors may be defined by researcher's interests, emergent research problems, disciplinary concentration, paradigms and paradigm shifts, as well as by the tools that mediate between the scholar and individual knowledge artifacts, and the aggregating structures (such as open access repositories) that organize individual knowledge artifacts into meaningful collections.



## 2.7 *OA implications for the scholarly publishing process*

The core supposition of OA in scholarly publishing is that by the practical implementation of the OA principles (electronic distribution of content, free and unrestricted access, author retains rights), the peer reviewed scholarly work will be disseminated as widely as possible and be freely accessible by researchers and other interested individuals alike.

In the cycle of scholarly communication, citation analysis has historically provided an important view of the impact of one article onto another. References play important role in authority building towards the acceptance of the article by the intended target audience. The importance of an article, and the ideas contained therein, is also reflected through the number of citations it has subsequently received (Borgman, 1989, p.590). Bibliometrics, as the study of relationships between texts through their citations and references, provides an insight into the formal channels of scholarly communication (p. 586).

A logical deduction is that citation analyses ought to provide the measures for the impact of OA articles on scholarly publishing process as well. For example, in a study that examines the self- archiving OA approach—pre-prints and post-prints on personal web pages, institutional and disciplinary repositories, professional sites, and course archives—using the ISI Web of Knowledge citation index, Antelman (2004) finds that "freely available articles do have a greater research impact. Shedding light on this category of open access reveals that scholars in diverse disciplines are both adopting open access practices and being rewarded for it" (p. 372). Antelman also finds that from the sample of OA articles across different disciplines she identified for her study,

mathematics has much higher percentage of articles made open access (69%), and philosophy has substantially lower percentage of (17%) (p. 375). There is also a correlation between citation and download hits of OA article. The more the paper is downloaded the more it is cited (see Antelman, p. 373). Harnad (May 2006), using RoMEO statistics shows that “Articles made ‘Open Access,’ (OA) by self-archiving them on the web are cited twice as much [than the non OA article], but only 15% of articles are being spontaneously self-archived” (Preamble, para. 1). Harnad also concludes that self-archiving has higher citation and hitcount impact when compared to articles in OA journals because it is supplementing the existing scholarly publishing systems. However, as a precondition to these positive OA impacts, the existing copyright models need to be restructured to provide researchers either full copyright to their work, or at least as much as necessary to enable researchers to deposit their work in institutional and disciplinary OA archives. This flexibility of leaving the totality of the copyright with the author will minimize the role of the traditional commercial publishers as “gatekeepers”, “managers” and barriers to creativity and innovation (Hunter, 2005, p. 9).

An investigation into open access and scholarly communication literature, addressing the concept of positive impact, reveals a set of bibliometric and citation analysis concepts and constructs adapted and modified to address the new context being constructed by the open access information environments. To understand the high-level discourse about the types of research articles related to OA research, 46 research articles were identified through Bailey’s Open Access Bibliography (2005), especially looking at various subsections of chapters titled Research Studies. Each study was tabulated (see

Appendix A) with its date of publication, the type of OA manifestation it addresses, and whether it is a perception / behavior study, citation impact study, a systems analysis and server log study, or usability study. This categorization of the research articles revealed that there was no apparent qualitative study in the mix exploring the new socio-technological dynamics in the knowledge production process by the introduction of OA repositories.

About twenty (20) of these studies address aspects of citation impact analysis, eighteen (18) discuss server deep log analysis for downloads and viewability, and eight (8) discuss usability studies. Twenty one (21) studies probe the perception and behavior of the researchers, librarians, tenure and promotion committees, financial management, copyright rules, etc., to assess and understand the shaping of the relevant actors due to the introduction of open access. What all these studies have in common is that they explore whether the new dynamics that has emerged out of the open access have indeed made scholarly articles more accessible.

However, a challenge has also emerged: across these research articles there is no clear consensus as to the meaning of article accessibility from operationalization perspective. Additionally, constructs such as article visibility, and discoverability have also emerged. For the most part these constructs are an attempt to adapt and appropriate constructs from the traditional citation impact studies and those that study the impact of web pages. Furthermore, researchers' information practices and the type of repository (disciplinary, institutional) may have additional implications in the quest to understand the interplay of OA in the scholarly publishing process.

## 2.8 *Defining article accessibility, discoverability and visibility*

Next, concepts and construct related to impact analysis are reviewed from the literature and three specific constructs are defined and adapted for this study based on their overlapping and common elements. These constructs are useful for this study as many of the participants responded to the interview questions with reference to these constructs when explaining the value of open access in their individual knowledge production as well as the role and value of open access for their discipline.

### 2.8.1 Article accessibility

Kling (2003, p. 594) defines article accessibility as follows: “Accessibility: Readers must be able to access the document in a stable manner over time. Libraries, publishers, and clearinghouses typically ensure accessibility by distributing and storing documents.” Accessibility is also related to the ability for an article to be clearly described and identified by a set of metadata (Kling & McKim, 1999, p. 898). Zhang (2001, pp. 630, 632) relates poor accessibility to the lack of standards for citing electronic sources due to the dynamic nature of electronic articles. There is not necessarily a cause-effect relationship between standards for citation of electronic resources (e-citation) and accessibility: for accessibility to be enhanced and standards for e-citation to be developed, stability over time is required and clear identifiable metadata standards for electronic articles are required. Chan (2004) suggests that the adoption of the OAI-MHP is related to article accessibility and visibility (Enabling technology, para. 3).

These previous definitions by Kling (2003), Kling and McKim (1999), Zhang (2001) and Chan (2004), imply the following about accessibility:

- a) Stable article availability over time, i.e., the content of the article can be read at any future time
- b) Article description with full set of metadata
- c) Ability to cite electronic articles, including standards for citing electronic article
- d) Interoperability between archives where open access articles are deposited for citation purposes as well as full content accessibility

*Operational definition of accessibility:* An article is accessible if its content can be read (by humans or machines). Note: Accessibility is an external property of the article, derived at least from the properties of the repository or other information environments where an article may be deposited and accessed from.

### 2.8.2 Article discoverability

Another construct is article discoverability. Antelman (2006) makes the following distinction between article discoverability and accessibility: “A publisher copy that the reader does not have access to, while discoverable, is not accessible and so not as useful as an open access copy. Similarly, an open access copy located on a website is not as discoverable as it could be if it were in a repository or a journal” (p. 93). This certainly suggests that accessibility and discoverability are separate constructs that are measured individually. McDonald (2006) investigates “the effect that discovery and access tools have on citation” (p. 41) by suggesting that access tools are implicated in the increased discoverability and accessibility of articles and increase article citation (p. 42).

Therefore, these previous definitions by Antelman (2006) and McDonald (2006) imply the following about discoverability:

- a) Articles can be discovered through the use of access tools that use either abstracting or indexing services based on metadata availability
- b) Discoverability as a construct is limited to the ability to know that such an article is out there, even though one might not have access to its content
- c) Discoverability is dependent on the information environment, i.e., specialized disciplinary archives are more supportive in discovering articles related to a research problem in comparison to institutional archives. Self-archiving in archives in general is more supportive in article discovery than personal Home Page self-archiving.

*Operational definition of discoverability:* An article is discoverable if its location can be determined. The article might or might not be accessible. Note: Discoverability as an external property of the article at least is derived from the properties of the access tools and the properties of the repositories or the information environment where the article is deposited.

### 2.8.3 Article visibility

Hedlund, Gustafsson, and Björk (2004) defined visibility in relation to subject-based indexes: “Subject-based indexes are a measure of visibility as they are very widely used by readers to identify or locate articles” (p. 205). Yue and Wilson (2004) define journal visibility as “the extent to which a journal is ‘seen’ by the academic community. It is gauged by the number of Abstracting and Indexing (A&I) services or databases that choose to include a particular journal” (p.320). Article visibility as a construct is congruent with the journal visibility in a sense that article visibility is derived from

journal visibility. In the context of open access disciplinary repositories, the visibility of an article can be derived both from its journal and as well as repository visibility.

Therefore, these previous definitions by Hedlund, Gustafsson, and Björk (2004) and Wilson (2004), imply the following about visibility:

- a) Visibility depends on article inclusion in subject based abstracting and indexing services
- b) Visibility depends on search engine indexing of archives; both general search engines and specialized search engines using the OAI-MHP

*Operational definition of visibility:* An article is visible if it is included in abstracting and indexing services, search engine indexes, as well as the inclusion in a repository. Before an article can be discovered, it needs to be visible. Note: Visibility as an external property of an article is at least derived from the properties of the access tools and the properties of the repository or the information environment where the article is deposited.

#### 2.8.4 The role of accessibility, discoverability and visibility

The attempt to appropriate, adapt and construct operational definitions for article visibility, discoverability, and accessibility in the context of OA repositories, reveals that the various interpretations of these concepts in the literature have used actors that are located in at least three different levels: a) metadata level (for description of articles, and access tools), b) interface level (links), and c) organizational level. Each digitally deposited article can be described by a set of metadata. Some metadata pertain to article's content, other to article's physical properties, article location in the form of URLs, as well as article's copyright properties, and ownership.

The exchange of the metadata occurs at the interface (technological) level, utilizing open standards and software for interoperability such as OAI-PMH. At the organizational level, articles are brought together into disciplinary or institutional repositories by various professional associations or institutional structures. All three levels can be analyzed as distinct contexts with specific technological and social actors and links, and performative capabilities.

#### 2.8.5 The role of CTAs for visibility, discoverability and accessibility

The summary of concepts related to visibility, accessibility and discoverability, suggest that all three depend on the availability of data (full text, abstracts, indexes, metadata, etc.), that for the most part are under the control of commercial publishers. Therefore, CTAs act as moderating gates to what is possible to be visible, discoverable, and accessible. And because the access tools that are used by the researchers to access the repositories also depend on the availability of the same data (full text, abstracts, indexes, metadata, etc.), it can be argued that the mechanism by which CTAs perform on the construct of visibility, discoverability, and accessibility is carried out through the aggregated role CTAs play in conjunction with the technical capabilities of the access tools—where the properties of the access tools are derived from their technological capabilities of interoperability and from the properties of the data they can access (that are mediated but the legal framework of the CTAs). This suggests that CTAs have performative capabilities that:

- a) Effect how researchers think about what can be made available in open access repositories



- b) Determine the capabilities of access tools with respect to the content they can act on (i.e., the article related content they can operate on), whether they are commercial publishers' portals or open access archives
- c) Link researchers to articles they need for research and scholarly endeavor; readers can have visibility to articles either directly or through the use of access tools.

The actor-network theory suggests that access tools as links and conduits have performative capability that can translate the commercial publishers' policies (the full continuum from very restrictive to very flexible) and inscribe their properties and dimensions into:

- a) Researchers' behavior to encourage them to engage in self-archiving, i.e., as enabler of OA
- b) Access tools' capabilities to take advantage of increased open content in repositories, and
- c) Researchers' behavior to use and cite articles available via open access and thus increase article circulation.

Thus, CTAs act both as barriers as well as as enablers of OA. They are barriers from the perspective that for articles to be accessible in OA repositories the respective CTA needs to allow for such inclusion in OA repositories, otherwise it will make the articles inaccessible in OA repositories. However, given the current status of proprietary publishing policies that are mostly restrictive, CTAs can enhance visibility, discoverability and accessibility of articles by allowing some level of article inclusion in OA repositories, as well as allowing the article related metadata to be shared and

accessed by access tools and other services, such as metadata harvesting and federated search services and portals.

## 2.9 *Critical actors as identified in the literature*

Based on the analysis of the literature presented in the Background section, the following is a list of actors that have emerged as critical for researchers' interaction with OA repositories in the scholarly communication process:

- Scholarly publishing cycle:
  - Knowledge production (Kling & Covi, 1995; Schauder, 1994)
  - Circulation (Kling & Callahan, 2003; Okerson, 2000)
  - Re-circulation (Kling & Callahan, 2003; Okerson, 2000)
- Disciplinary differences implications on scholarly production:
  - Sciences (Palmer, 2005; Fry, 2006; Bates, 1994; Whitley, 2000)
  - Humanities (Palmer, 2005; Fry, 2006; Bates, 1994; Whitley, 2000)
- Access tools as enablers and inhibitors:
  - OA standards (OAI; OAI-PMH; Coleman & Roback, 2005; Suleman & Fox, 2002)
  - Indexes, abstracts, references, etc. (Harnad, 2006; Borgman, 1989; Antelman, 2004)
  - Copyright Transfer Agreements – CTA (Beier and Tschida, 2003; Coleman and Roback, 2005)
  - Access resources (Palmer, 2005)
- Article level of openness:
  - Visibility (Hedlund, Gustafsson, & Björk, 2004; Yue & Wilson, 2004)
  - Discoverability (Antelman, 2006; McDonald, 2006)
  - Accessibility (Kling 2003; Kling & McKim, 1999; Zhang, 2001; Chan, 2004)
- Knowledge artifacts
  - Articles (Van de Sompel et al., 2004; Schauder, 1994 )

- Repositories (Van de Sompel et al., 2004; Borgman, 1999)
- Types of OA repositories (BOAI, 2002; Hunter, 2005; Willinsky, 2005):
  - Institutional
  - Disciplinary
- Open Access (BOAI, 2002; Hunter, 2005; Willinsky, 2005)

These are only the initial sets of actors that have emerged as important from the literature review. As a result of this study, however other important actors will emerge.

### *2.10 Key terms*

These key terms and the associated definitions emerged from the literature review. Some have been clarified, combined, or otherwise simplified, and adapted for specific use in this study.

Actor – in the context of scholarly communication, actors are entities that can be human or non-human such as researchers, access tools, journals, articles, books, computers, social structures, information systems, repositories, tenure committees, etc.

Article accessibility – An article is accessible if its content can be read (by humans or machines).

Article discoverability – An article is discoverable if its location can be determined. The article may or may not be accessible.

Article visibility – An article is visible if it is included in abstracting and indexing services, search engine indexes, as well as the inclusion in a repository. Before an article can be discovered, it needs to be visible.

Disciplinary knowledge network (knowledge ecosystem) – a network of knowledge artifacts available to the participants of a scholarly discipline through which and by which the epistemic elements of the discipline are defined.

Disciplinary knowledge production context - a context comprised of disciplinary norms and cultures, including the accepted information practiced and accepted disciplinary knowledge networks in a specific discipline.

Individual knowledge network – a localized knowledge network, defined by a subset of artifacts from the disciplinary knowledge network used by a scholar in the production of a specific scholarly work.

Individual knowledge production context - a context comprised of the local and immediate processes and localized knowledge networks used by an individual researcher.

Inscription – an ANT semantic construct, denoting a process by which actors perform on other actors shaping their properties and dimensions.

Link/relationships – a conduit that associates two actors.

Lived experiences and perceptions layer – is manifested through the themes of discourse based on the interviews with the researchers reflecting on their experiences as they interact with the open access repositories.

Network topology – a complex entity constructed by two or more actors connected between themselves through various links or communication channels.

Open Access (OA) – the free and unrestricted access to scholarly literature (pre-prints and post-prints) through the world-wide networked electronic distribution.

Organizational layer of open access repositories and access tools – is manifested as arXiv and PhilSci, with properties instantiated by the organizing and management structures of the open access repositories and the access tools and the enacted policies that govern their operation.

Repository – an organized digital collection of articles (pre-prints and post-prints) with capabilities for researchers to deposit their work and also be able to discover scholarly work that has been deposited by others.

- Institutional repository – a repository organized, established and maintained for the purpose of a specific institution to enable its researchers as authors to deposit their scholarly work, and read the work of others.
- Disciplinary repository – a repository organized, established and maintained for the purpose of specific disciplinary field to enable the researchers as authors that associate themselves with the discipline to deposit their scholarly work, and read the work of others.

Researchers – For this study, the participants are individuals that are affiliated with scholarly disciplines and produce scholarly works.

Technological layer of open access repositories and access tools – is manifested as the software that enables the open access repositories (such as EPrints), with features and functional capabilities as built by the software developers.

Translation – an ANT semantic construct, denoting a process that enables the realignment of actors and their relationships by making them equivalent.

### **Chapter 3.           The Research Problem**

The literature review, as presented in the Background and literature review chapter, suggests that the loosening of copyright rules by commercial publishers to allow for pre-prints and post-prints of articles to be deposited in archives outside of their control, and the availability of access tools and OA repositories to researchers, should realign and restructure the relationships in the ecosystems of scholarly publishing and potentially bring forth new links as a result of the new phenomenon. Some of the new and realigned links have resulted from the transformation of print publishing by networked scholarly publishing (supported by the emergence of the Internet)—that has enabled the emergence of a multitude of access tools and organizing structures for the management and organization of electronic scholarly artifacts. This in turn has enabled researchers to use their time more efficiently by moving portions of their knowledge production process away from print by using electronic publishing capabilities. Understanding the dynamics of these new links and associations in the context of OA repositories as an emergent set of actors in the knowledge production process requires an understanding of researchers' use of self-archived articles and other knowledge artifacts from OA repositories, where OA repositories are the conduits that link researchers amongst themselves (as producers of knowledge and as readers) and they also link researchers to the institutional or disciplinary context where the repositories are embedded. The emergence of these new dynamics in the knowledge production process and the scholarly communication more broadly by the introduction of OA repositories can be understood and described through the new set of actors that appear to be important, but

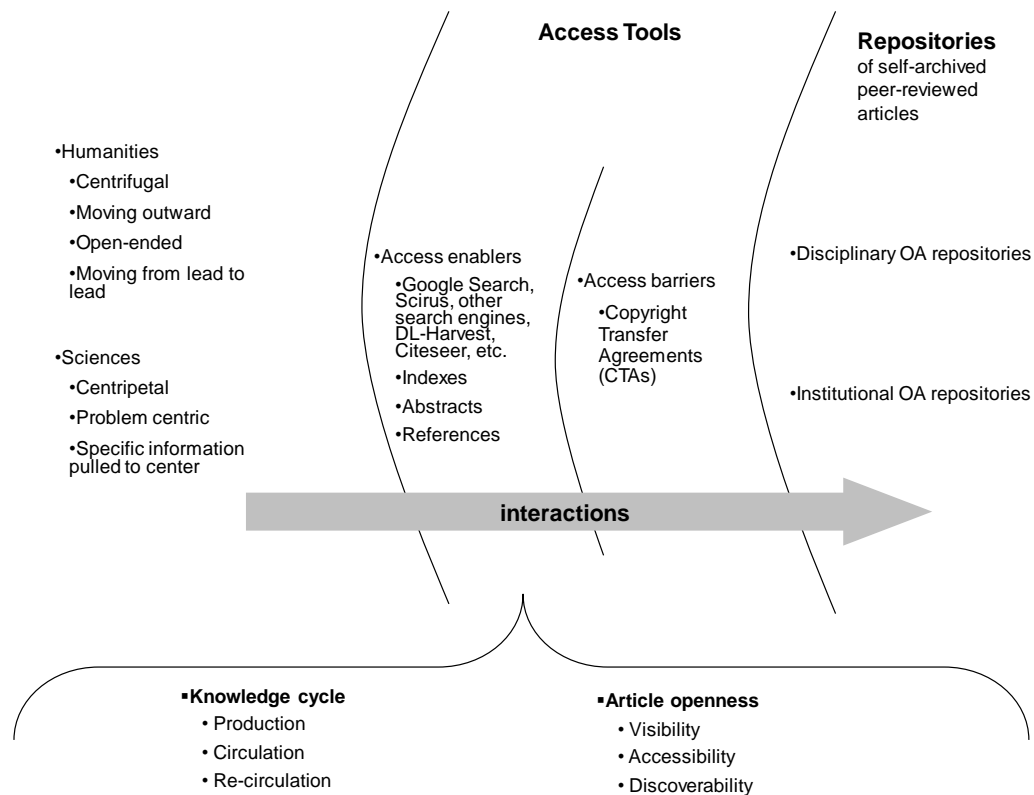
there does not seem to be an explication about how are they relevant, or the nature and dynamic of the interactions between the actors.

Further, the analysis of the 46 research articles related to OA, as it has been mentioned earlier, reveals that qualitative research studies are lacking in the overall research about OA (see Appendix A). Yet, there is no apparent systematic methodological approach about how to understand and impartially describe the interaction dynamics between researchers and open access repositories that includes both technological and social elements. Therefore, this study will elucidate the implication and the role of the socio-technological dynamics of open access repositories in researchers' knowledge production process as embedded in the broader frameworks of scholarly communication and individual and disciplinary knowledge production contexts.

To further structure and frame the research problem and the research questions, the conceptual map from Miles and Huberman (1994, p. 18) has been adapted as a visual aid. Presented in Figure 3 is researchers' interaction with different types of repositories as mediated by different types of access tools and the different types of CTAs. Using the closed rational and open natural systems models view suggested by Covi and Kling (1996), one way to look at Figure 3 is to think of the boundaries as gates: the more open these gates, the better the accessibility to the knowledge artifacts in repositories. In other words, from the open access perspective, for the content of an article to be visible, discoverable and accessible without subscription, the copyright transfer agreement should not prohibit such discoverability and accessibility. However, removing the copyright barrier is not sufficient; this only shows the potentiality of the article to be

discovered and accessed. For articles to have an impact on scholarly communication, researchers should have an easy way of discovering articles of interest in their scholarly information seeking process. In addition to enabling unrestricted access to peer-reviewed articles, OA repositories enable the meaningful organization of other knowledge artifacts such as pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc., and thus increase their use through increased visibility, discoverability and accessibility.

Figure 3: A conceptual framework of researchers' information access modes, access tools, and repositories



Note: This figure that has been constructed based on the understanding from the literature review.

Thus, the intent of this study is to provide an understanding of the implication of open access repositories on researchers' knowledge production process by understanding



and explicating researchers' lived experiences and perceptions as they use the open access repositories. In the context of open access in scholarly publishing, the literature suggests that researchers as authors, their access tools and the repositories are relevant elements in the knowledge production process, and that researchers' interaction with repositories is indeed a composite construct of at least two sets of interactions: a) the interaction between the scholar and the access tools, and b) the interaction of access tools with repositories. Further, researchers' knowledge production processes are enacted within the disciplinary knowledge production context, where researchers' information practices inform and are also informed by the broader context.

### *3.1 Research Questions*

The main intent and goal of this study is to understand the implication of the OA repositories on researchers' knowledge production. Recognizing that researchers' knowledge production process is situated within the broader disciplinary and institutional socio-technological and cultural context, through the activities of the researchers, the research questions will elucidate the implication of OA on researchers' individual knowledge production context across the lived experience, the organizational and technological layers, and also the implication for the disciplinary knowledge production context. The implications are compared and contrasted among and between the arXiv scientists and the PhilSci scholars, as representatives of different approaches to scholarly information practices and distinct epistemic cultures. Based on this intent and goal, the research questions are:

RQ1: How do researchers experience and perceive the role and value that OA repositories and access tool provide in their knowledge production process?

RQ2: How are the properties of the organizational and technological layers of OA repositories and access tools implicated in researchers' individual knowledge production contexts?

RQ3: How are the actors and properties of the broader socio-cultural and technological context, including disciplinary norms and cultures, implicated in researchers' knowledge production context?

RQ4: How are the performative agencies of the key actors and contexts implicated in researchers' individual knowledge production context?

RQ5: How are the organizational and technological layer properties of the open access repositories and the access tools related and associated?

RQ6: How are the open access repositories and access tools implicated in the structuring of the knowledge production contexts of the arXiv scientists and the PhilSci scholars?

#### **Chapter 4. Theoretical framework: Actor-Network Theory**

The technological deterministic and the social deterministic views of scholarly communication when applied in isolation from each other provide partial and skewed explanation about the socio-technological changes related to scholarly publishing. The theoretical approach used in this study to explain how researchers use self-archived open access repositories is informed by the actor-network theory and methodology. The socio-technological co-constructionist approach to studying information works emerges as the most congruent approach to explore and describe constructs that exhibit both technological and social manifestations. The actor-network theory aims to describe and explain in detail how various human and non-human entities affect each other in a complex web of interconnections: “We are emphasizing this process of mutual shaping because it is important to understand that actors are not simply shaped by the networks in which they are located (although this is certainly true), but they also influence the actors with which they interact” (Law & Callon, 1997, p. 25).

Buckland (1991) similarly argues for the holistic approach within the context of information environments and digital libraries: “Since information and information handling is pervasive in human activities, an exploration of information systems that did not include the social, economic, and political context and the broad social role of information will be seriously incomplete” (p.9). Electronic scholarly publishing with its inclusive and tangential actors, both social and technological, is a relatively stable socio-technological system bound together with a set of conduits and associations that link together the various social and technological actors.

Approaching the scholarly publishing process from ANT perspective has theoretical and methodological implications at three different levels:

- a) As a high level meta-theoretical thinking:
  - ANT suggests the co-constructionist approach because all actors, whether social or technological, are associated and linked to each other.
  - ANT empowers the investigator with a conceptual framework and terminology to develop network of social and technological actors impartially.
- b) As a way to knowing where to look and how to look in building a socio-technological network and especially identifying associations and links and tracing them.
- c) As a way to analyze the technological and the social together, instead of analyzing the social separately from the technological because such analysis eventually involves reductionism and comparison of “sanitized” results where details are lost. With the technological and social being analyzed together at the very detailed level, a more granular view of the interactions will emerge.

#### *4.1 The overall theoretical orientation – co-constructionist approach as the meta-theoretical direction for the study*

The overall approach to this study is informed by the actor-network theory and the associated methodology. Creswell (2003) suggests that this paradigmatic approach is a type of overarching theoretical lens that “provides a framework for topics of interest, methods for collecting data, and outcomes or changes anticipated by the study” (p. 16). However, with actor-network it is more appropriate to think as a prism instead, because

in addition to the new and novel approach, it also provides a methodological approach about how to understand actors and associations through their constitutive elements, identifying their properties and dimensions and explore them at the appropriate level for the problem at hand.

Bijker and Law (1997) have shown that technologies are not produced in isolation from social structures (p. 3) and that technologies affect the social structures within which they are embedded (p. 4). The actor-network theory postulates that actors affect and modify each other by the very fact that they are linked: "... entities [actors] take their form and acquire their attributes as a result of their relation with other entities" (Law, 1999, p. 3). This interplay between actors and associations in the context of OA scholarly communication process lends itself as a potential candidate for description and understanding by the actor-network theory and its methodological framework (Law & Hassard, 1999).

#### *4.2 Methodological implications of socio-technological analysis*

The language and terminology provided by the actor-network theory and methodology emerge as enabling tools to discuss and analyze things of the socio-technological nature simultaneously and impartially in one framework. Thus, the actor-network theory aims at remedying the methodological challenge that emerges from the difficulty in trying to make sense of findings by studying the technological aspects separately from the social aspects.

Considering the co-constructionist and socio-technological nature of the open access phenomenon related to scholarly publishing (encompassing researchers, readers, copyright transfer agreements, indexing and abstracting services, information

environments, access tools, protocols, standards, tenure and promotion committees, etc), studying the technological and the social factors of the open access phenomenon as one network is an extension of their continuity in naturally-occurring settings in which the knowledge producing interactions take place across technological and social dimensions. As it has been already mentioned, Bijker and Law (1997) have demonstrated that technologies are not produced in isolation, rather they are amalgamates of their intrinsic properties and a reflection of their non-technological context, and further elaborate and describe how technologies modify and define the social structures within which they are embedded.

Demonstrating with Figure 2, to analyze the social and the technological together would mean to remove the distinction between technological and social layers altogether and consider them to belong to the same layer and that they are heavily interconnected with the ability to inform and shape each other.

#### *4.3 Consideration of alternative theories and frameworks*

DeSanctis and Poole (1994) suggest the Adaptive Structuration Theory (AST) as a mechanism to examine the change process in a given context by looking at the type of structures provided by advanced technologies (inherent structures), and the structures that actually emerge in human actions as people interact with these technologies (p. 121). AST is more specific to the interplay between the structures provided by advanced information technologies and social structures within which they are embedded. Its emphasis is on how the features and properties of an information system are acquired, appropriated and constructed, and on how these specific structures affect the organizational structures.

In the onset it would appear that actor-network and AST attempt to explain the same: "So, there are structures in technology, on the one hand, and structures in actions, on the other. The two are continually intertwined; there is a recursive relationship between the technology and action, each iteratively shaping the other" (DeSanctis & Poole, 1994, p. 125). However, they seem to complement each other in a sense that actor-network provides us with the language to understand how actors affect each other, while AST is mostly concerned with the aspects of how a particular actor (information system, social structure, etc) is constructed and modified in the iterative process: "New social structures emerge in group interaction as the rules and resources of an AIT [advanced information technologies] are appropriated in a given context and then reproduced in group interaction over time" (DeSanctis & Poole, p. 129). AST strongly emphasizes the process of appropriation of information technologies in a given setting, suggesting that information technologies in their initial implementation come with pre-defined feature (i.e., the spirit) that are being appropriated by the users. The appropriation in turn has effect on the construction and modification of the social structures that in turn redefine how the information technology is used, often in ways not intended by the designers. Thus, it appears that both actor-network and AST are implicitly informed by systems theory (Churchman, 1968; Ruben, 2000), complementing each other by providing a language and concepts that are necessary to understand the interactions and the mutual shaping of human and non-human entities in various environments.

Socio-Technical Interaction Network (STIN) is another candidate framework that can explicate the socio-technological aspect of a context. This approach, based on the actor-network theory, has been developed by Kling, McKim and King (2003):

A Socio-Technical Interaction Network (STIN) is a network that includes people (including organizations), equipment, data, diverse resources (money, skill, status), documents and messages, legal arrangements and enforcement mechanisms, and resource flows. The elements of a STIN are heterogeneous. The network relationships between these elements include: social, economic, and political interactions. (p. 48)

From the perspective of operationalization of a methodology in order to move ANT from theory to practical use for specific problem, STIN appears as an attempt to operationalize ANT in the context of collaborative environments supported by technological artifacts. To help in this operationalization, the authors provide a guide about how to specifically model a socio-technological context as a STIN (p. 57). The resemblance of STIN to the Actor-Network Theory (ANT) is not accidental (see Latour, 2005, Law & Hassard, 1999). Kling, McKim and King state that “STIN models differ significantly from actor–network theories in that STIN methodology requires that the analyst make attempts to understand the characteristics and scope of the interactors [i.e., associations/links] ahead of time, rather than taking the interactors as they come and following them through use or development of the forum [i.e., information/communication environment]” (p. 57). While ANT provides a framework about how to do an ethnomethodological study by following the actors and more importantly actors’ associations in their socio-technological context as they occur, STIN suggests that the task of identifying the actors and their associations should be performed ahead of time. There certainly is a challenge with identifying actors and interactors ahead of time. By doing so, we will be imposing our knowledge frame and



understanding on the environment, as a prism, that might exclude us from seeing and identifying critical and relevant actors and associations in emergent fashion. Reading closely Appendix B (Research Methods) in Kling, McKim and King (2003) it may appear that STIN is no different than actor-network in its methodological approach—even with STIN the authors had to go back to re-interview the informants, and then introduce new associations into their model. If there is a substantial difference between STIN's and ANT's methodology, it is in the fact that STIN is an appropriation of ANT for the scholarly communication contexts and provides some guidelines for new researchers as to what interactors or associations between the actors are relevant. As it relates to this study both approaches have acted as a guide to identify the importance of access tools as associations and conduits that link authors and readers with the knowledge artifacts deposited in the repositories.

ANT is more congruent approach for this study, as it provides an overall framework that guides this study, provides a language with which to impartially describe and analyze human and non-human actors, and it also enables the tracing of actors via their relationships to the local and immediate networks as well as the global network and the contexts they belong to.

## **Chapter 5. Methodological perspectives**

The following section examines the methodological frameworks used in this study. ANT, grounded theory and phenomenology are described in context and a rationale is given for their choice and combination. This section also describes the choice of methods that correspond to the respective frameworks and approaches.

### *5.1 Grounded theory and phenomenology*

Given the new phenomenon, i.e., researchers' interaction with access tools and OA repositories, there is congruency with the phenomenological approach that enables understanding researchers' lived experiences as described in their own words (Creswell, 1998, p. 55), with the main goal to help with the understanding that a single unified meaning of experience or perception of a phenomenon in question does exist (p. 55). Grounded theory complements the phenomenological approach by providing a systematic set of methods to analyze the data gathered from interviews and documents, identifying any underlying common structures and processes enacted by all researchers as they experience their interaction with OA repositories.

Once the concepts, categories, properties and dimensions are identified using grounded theory approach, a set of contexts will emerge. One network topology may represent the knowledge production process in the context of OA repositories. A knowledge artifact that is produced is a form of a network topology in the sense that it represents a collection of data from various resources that are logically associated to represent a new knowledge construct. In a sense, network topologies or context are meaningful constructs and models, but also graphical topological representations that

make visible complex sets of socio-technological links and associations that are bound together in a dynamic but stable network topology, enacted by researchers' information practices as they seek articles to create new knowledge.

After the concepts, categories and their properties and dimensions are identified, the semantic elements of ANT are applied as an interpretive narrative to link and associate the concepts and categories across their properties and dimensions.

## 5.2 *Semantic Elements of the Actor-Network Theory*

This section describes the terminology enacted by the actor-network theory and methodology. The semantic elements of ANT empower the investigator to discuss through constructs that by themselves are composite entities as described in the definitions in this section. This terminology is neutral to the technology and social structure duality, thus it is used to describe social and technological actors impartially as they come together to develop and enact socio-technological contexts (network topologies).

*Actor* is an entity, either human or non-human that is a constitutive element of a network topology. In the context of scholarly communications, actors can be human or non-human such as researchers, journals, articles, computers, social structures, information systems, information environments, etc. All these actors and their relationships dependent on their positions in a pertinent network topology, and therefore are constantly in flux, yet relatively stable. For example, an article can be an actor in the network topology bounded by the journal in which it is published. The same article can also be part of the network topology bounded by the OA repository in which it is being deposited. It is also an actor in the network topology defined by a disciplinary field.

Unless a specific article is an extreme case that defines the discipline in which case it will be strongly linked to other similar articles, most articles have stronger relationship in the network topology bounded by the journal construct, and weaker relationships in the network topology bounded by the discipline.

*Actor-network, or network*, is a complex socio-technological context, a network topology of logically grouped entities associated and linked to each other via some relationships. For example, a single interaction can be described as a network that is embedded in the larger context (i.e., the network that describes the knowledge production process that among other contains multiple interactions). Also, a disciplinary OA repository is a network topology of articles, pre-prints and other knowledge artifacts for a specific discipline or field of study that defines the disciplinary repository. An institutional OA repository on the other side is a network topology of knowledge artifacts produced by the researchers and other members of the institution that built and maintains the repository.

*Association/Link* is what connects two actors. For any actor to be a meaningful entity and matter within a network structure, it must be somehow connected and linked to other actors in its near vicinity. Depending on the perspective and consideration of functionalities of a given network topology, a particular element could be treated either as an actor or as a link. For example, a co-authored article is an actor in the journal it is published, but it is an association that links two researchers. The associations and links in institutional OA repositories are manifested as topical collections, department collections, etc—and they are external to the artifacts themselves. For disciplinary OA

repositories, the associations that link the knowledge artifacts together are disciplinary related and are internal (i.e., content) to the knowledge artifacts themselves.

*Inscription* is the act or process by which actors perform on other actors, shaping and defining their properties and dimensions as co-construction occurs. The properties and dimensions of any particular actor or network in part are acquired a result of a complex inscription process by human and non-human actors. Human actors are able to inscribe onto non-human actors, as well as non-human actors are able to inscribe onto human actors (Akrich & Latour, 1997). An example of how humans inscribe into OA repository would be that the features and capabilities of the repository are designed to closely match the information practices of researchers, i.e., making the repositories easy to use by asking the researchers/users to help with the design. On the other side, if the researchers have to modify and change their information seeking tasks to the features and tasks that are embedded in the repositories, this would be an example of the repository inscribing its properties onto the researchers information practices. The inscription process is conditioned upon how open the actors or the networks are, and the congruency between the actors and the associations amongst them. Thus, the ability to inscribe is dependent on the openness of actors' boundaries and actors' mutual congruency. An example of inscription will be to achieve the desired results around article distribution as defined in a specific CTA. If a CTA states that the article cannot be deposited in OA repository, a successful inscription will be considered if the article is not deposited in OA repository, thus achieving its intended outcome. The congruency in this case will be achieved by the fact that the researchers as agents that can act on the CTAs actually know what they mean and are able to either follow the requirements of

the CTA or not. Inscriptions are not always one-to-one however. Whether an article makes it into an OA repository is also mediated by other actors who are trying to inscribe their properties onto the article. This means that more than one actor is trying to inscribe their properties onto the article for the relevant topology defined by the OA repository. Another way the concept of inscription is used in this study is to explain the performative capabilities of actors and how they translate into the other actors, or how the properties and dimensions of multiple actors are bound together to enact another construct/actor. An example of this would be the complex network topology named repository. A repository is an amalgam of software, standards, metadata, processes, workflows, interfaces, interoperability policies, organizational structures, etc. that are bound together to form a new entity called repository.

*Openness.* The concept of openness is another important concept of the actor-network. Without openness there cannot be any communication between actors in a given network. Therefore the boundary of any given actor (or network) is porous as the enabler of exchange, and the nature of the openness is related to the level of exchange that is possible between the inside and the outside: "... a distinction is made between inside and outside [of an actor] and a set of exchanges between the two is defined and regulated ... those who are outside find themselves compelled to participate in those exchanges" (Law & Bijker, 1997, p. 294). In the case with OA repositories, various access tools act as mediators and partially determine the level of openness of the repositories based on the interoperability between access tools and the open access repositories.

*Translation.* The *translation* process of the actor-network methodology has emerged as one of the key concepts in the attempt to explain the potential performative influence between actors. The translation process, described by Law (1999) as “... the process or the work of making two things that are not the same, equivalent” (p. 8), suggests that properties and dimensions are transferable from one actor onto another and are carriers for the transformation process, subject to the congruency between the two things. An example of this would be the open access policies that may be set for a specific institutional repository. The institutional repository would be open access to the degree that the contributors follow the policy. In this case, the policy that promotes open access is the agent/carrier of change, reflecting what was inscribed into itself by other actors. However, a close look at the definition also reveals another important sense in which the concept of translation has been used, especially in Law’s language that the term “actor-network theory” was transformed into the “ANT” acronym. In this case, a complex set of theoretical and methodological actors was replaced and labeled (read as aggregated into) as ANT. From this, it is evident that the definitions of actor, network and link are relative with respect to each other, depending on the point of reference. For example, a scholarly article is an actor in one volume of a journal. It is however also an actor in the discipline that claims the particular journal as a representative of its field of study. At the same time, the journal itself is an actor amongst other journals in the same field of study, as well as in the OA repositories that it is deposited. Thus, if reductionism (i.e., translation) is attempted one may define all elements of the actor-network as special cases of the definition of an actor depending on the element’s intrinsic and extrinsic properties and dimensions. As the translation process is enacted the actors and

relationships of a specific network topology are realigned and repositioned with respect to each other and the broader context. Further, as translations are enacted they can also result in the emergence of new structures.

The *properties and dimensions* of actors can potentially translate the corresponding properties and corresponding dimensions of the actors they have acted upon, subject to the degree of openness and congruency of the links and the actors. An emphasis in this process is transferability of properties and dimensions from one actor to another through translation. The process of translation as described by Law (1999) suggests that the “things” that can be made equivalent contain corresponding or congruent properties and dimensions. Like the actors in a network topology, the relations (i.e., links) have properties through which the actors can potentially perform on the rest of the network and be performed by it (pp. 6-7). As with the actors’ properties and dimensions, link’s properties and dimensions ought to be congruent and corresponding for any translation to occur. The key however is that the network actors can be human or non-human, with both human and non-human actors able to act upon each other and induce translation. For example, if the properties of a CTA are described as enablers for OA (i.e., if publishers give permission for the articles to be deposited in OA repositories), the specific set of articles that are subject to the CTA should reflect these properties, i.e., should be reflected in the articles being deposited in OA repositories.

*Boundary:* An important characteristic of any network is its boundary. Defining the boundary is important for the successful utilization of the actor-network methodology. The boundaries of a given network and the relationships amongst its constitutive entities with their scope of influence are identified with the construct *topology*. An actor-



network topology is usually described as logically grouped entities or elements associated and linked to each other via some relationships (defined via properties and their dimensions). An actor can belong in multiple topologies simultaneously, performing and behaving differently in various topologies depending on its relative position in the respective network. For example, one boundary for a specific article can be an OA repository, a journal, a topical collection, or a personal Home Page.

Furthermore, inscription's and translation's ability to act at distance across time and space is an important capability postulated by the actor-network framework. It is the translation and realignment carried via actor's inscriptions that enable the actor to transfer its properties and dimensions to other actors in its immediate topologies (at various levels), subject to actors', links' and topologies' degree of openness. The processes of inscription and translations are in constant flux and iterative in nature thus enabling relative stability in the most localized actor-network. The perceived stability is actually performative in nature. It is relative stability, enabling entities in any given network to maintain themselves within the boundaries defined by the immediate network.

Latour (2005) further describes and distinguishes actor-network theory (ANT) as an epistemological tool – a different framework for looking at the world around us. While the traditional sociology studies the “socio” of actors as separate experience from the technological, ANT is meant to provide an investigator with the ability to understand and explain the associations (or links, or conduits) between actors with the ability to carry on translations among actors.

### 5.3 *Links and associations in building network topologies*

One of the key tools provided by the actor-network methodological approach is the ability to know where and how to look in a socio-technological context of investigation, the ability to trace and identify the entities and interactions dynamics that are critical to the specific network topology, and learn from the entities and their associations how and why they come together to build the actor-network, without imposing our a priori understanding:

For us, ANT was simply another way at being faithful to the insight of ethnomethodology: actors know what they do and we have to learn from them not only what they do, but how and why they do it. It is us, the social scientists, who lack the knowledge of what they do, and not they who are missing the explanation of why they are unwittingly manipulated by forces exterior to themselves and known to the social scientist's powerful gaze and methods. (Latour, 1999, p. 19)

Far from being a theory of the social or even worse an explanation of what makes society exert pressure on actors, it always was, and this from its inception, a crude method to learn from the actors without imposing on them an *a priori* definition of their world-building capacities. (p. 20)

It is this process of actor and association understanding, by tracing and reaching to the next linked actor and association that assist with building the network topology of the elements that define the boundaries of a network.

*The tracing process through an example.* As a methodological approach, tracing the links to identify how and why a specific actor is associated with other actors, in the

context of knowledge production, the process can start by scrutinizing what is believed to be an important actor for the given context. If we start with a specific access tool, we can learn about its properties and its technological construction from document analysis. We can learn about its situatedness and use by interviewing the researchers. Understanding the technological construction will reveal other associated actors such as metadata and protocol standards. Enacting standards by themselves is yet another context with its own set of actors and links (e.g. competition between competitive metadata standards for acceptance by the broader community). The same process can be repeated for the interoperability protocol as well. Tracing the metadata and protocol standards might further link to the open source mode of software development and its competition with the proprietary modes of software development. Tracing the links related to the access tools' situatedness and their use by the researchers will require interviews with researchers. From the interviews we can learn that the same access tool is used by researchers in various disciplines; it might be used to access already discovered articles, or to discover articles related to the topic of interest for the creation of a specific article. The articles themselves can be used to discover new facts or relevant data, or to confirm an existing argument pursued by the researcher.

The question arises as to how far the links need to be traced. Here comes the notion of boundary that the investigator might roughly be aware beforehand. The extent to which certain actors will be determined to be critical or not will depend on how strongly they are able to inscribe their properties through translation onto other actors through associations, i.e., to what extent do they determine the nature of the network topology.

For this study, the starting boundary that defines the network topology is defined by the researchers, access tools (including CTAs as potential enablers and barriers), and the repositories, in the process of knowledge production. What links them together are researchers' actions to discover and access articles in a repository for the purpose of using the articles in a knowledge artifact. After building the extended network topology through association tracing, it might emerge that metadata standards as a source for organizing structures in the repositories are very critical for the type of tools used by specific researchers. If however the access tools support different metadata standards inclusively, the metadata might not be critical for the choice of access tools, but specific metadata standards might be very critical in determining the description and organizational aspects of articles in a repository—partially due to the extent of what the metadata make possible to be described.

## **Chapter 6.       Methods and methodological considerations**

This section describes the congruency between the various steps of the research process, presents the rationale for the choice of methodological approach, and it details the research design and the steps for conducting this study.

### *6.1 Research process congruency*

This sub-section describes the logical congruency of the various steps in the research process within the context of qualitative approach.

As it has been described earlier in the Research Questions section, the intent of this study is to understand and describe the implications of open access repositories on researchers' knowledge production process. In essence, to understand and describe an emergent phenomenon within the context of scholarly communication, a qualitative approach is most congruent because it enables to understand and learn what are the relevant actors and relationships and their corresponding properties that matter for the newly instantiated phenomenon and the related context. The qualitative approach enables the discovery of emergent concepts, categories, themes and patterns inductively from large amount of data collected from the participants in similar settings that experience the same phenomenon (Creswell, 2003, p. 131-35).

Talja, Keso, and Pietilainen (1999) emphasize that the setting or the context should be accounted for when investigating researchers' lived experiences and meanings. The context where lived experiences are grounded in should not be objectified and trivialized; rather, the context should be accounted as an interactive and intertwined set

of actors performing in the instantiation of the phenomenon itself: “From a metatheoretical viewpoint, context is the site where a phenomenon is constituted as a research object” (p. 751). The context is not objective in the sense that it is out there to be analyzed, rather, the context within which the experienced phenomenon is enacted is co-constructed and instantiated. Thus, specifically for this study, borrowing from Talja, Keso, and Pietilainen (1999), researchers’ individual knowledge production contexts need to be understood as part of the broader context where researchers instantiate their search process to discover and access knowledge artifacts in open access repositories. Approaching the context as a dynamic topology of actors that is being co-constructed as researchers use open access repositories, is congruent with the co-constructionist approach of this study, where ANT as a meta-theoretical framework provides the higher-level context that enables a number of contexts to be brought together.

For this study, a closer examination of the actors identified in Figure 2 shows that the actors broadly belong in three different contexts: a) researchers’ perceptions of their lived experiences as they access the open access repositories, b) the technological context as manifested through the numerous technologies that are used as the researchers interact with the repositories, and c) the organizational context as manifested through the organizing and management structures that have made the various technologies operational and usable for specific purpose.

Researchers’ interactions with the OA repositories are instantiated as amalgamates with elements from these three contexts: lived experiences, technological and organizational contexts. The interaction is a process driven activity of specific tasks and actions, thus, there is a congruency with data collection that emphasizes on lived

experiences and understanding the interaction process. Given the varying nature of the preliminary actors identified in the Background and literature review section as critical in understanding this phenomenon, and the goals set to identify the concepts, categories, themes and understand any systematic patterns in researchers' interaction with repositories, grounded theory methods for data coding, analysis and interpretation emerge as congruent methods for this study.

ANT's semantic elements, such as inscription and translation, are congruent with the interpretation and discussion of findings in socio-technological co-constructionist context, as both inscription and translation are neutral to the social and technological duality and can help with the interpretation of how actors and relationships can perform on each other and realign their positions and roles with a network topology. These semantic elements enable the understanding and description of the links and actors in terms of their properties and dimensions. Translation and inscription will help with understanding of how properties from one actor (or broader context) are translated and inscribed onto other actors (or categories). Because ANT enables the understanding and description of socio-technological links of a phenomenon and its context, it will help explain how local and also seemingly distant and invisible socio-technological actors realign researchers' knowledge production processes and knowledge networks.

In other words, researchers' lived experiences provide an insight into how the researchers locally experience their interaction with open access repositories in the context of knowledge production, but their lived experience is also informed by the broader scholarly context (i.e., disciplinary norms and cultures). For example, the researchers do not directly experience metadata exchange between access tools and open

access repositories, nor do the researchers experience the organizing structures (for example board members and management) of a specific repository. However, the metadata and the board members seem to play a role in how the researchers experience their interaction with the open access repositories, even though the researchers might not be aware of such links. The translation and inscription enable the understanding and description of the relationships between lived experiences and the metadata and the board members, by describing how are they linked, associated, and the nature of the performative abilities and realignments instigated by the relationship.

Thus, as researchers' interactions with open access repositories traverse across socio-technological and organizational contexts, in addition to understanding researchers' lived experiences, the technological and organizational aspects of the access tools and the open access repositories need to be taken into consideration for more complete understanding.

## 6.2 *Research design*

Miles and Huberman (1994, p. 30) suggest that sampling and boundary setting should happen by considering four parameters: setting, actors, events, and process. For this study, the setting is defined by the OA repositories that are accessed by the researchers with the intent to use the knowledge artifacts contained therein in their knowledge work. The process is researchers' information seeking, i.e., searching, discovering and accessing knowledge artifacts in OA repositories. The event is scholar's production of knowledge artifacts. The initial actors are the researchers, access tools and OA repositories.



The boundaries of the setting are defined by the combination of researchers' information practices and the repositories being accessed. In order to account for the two types of researchers' access modes and two types of potential repositories, the cross-section possibilities result in the following combinations:

- Humanities and sciences disciplinary repositories
- Institutional repositories in a predominantly humanities research university and institutional repositories in a predominantly sciences research university.

While disciplinarily open access repositories are being actively used by researchers in their knowledge production, institutional repositories are predominantly organized and enabled by specific institutions with the intent to collect and preserve the output of their specific research community. Davis and Connolly (2007) in their explanation of why the institutional repository at Cornell University is not used by the faculty as intended enumerate a few concerns; including that faculty rely mostly on other means to access scholarly work pertinent to their disciplines. Most prominent in complementing the traditional scholarly process are the disciplinary repositories and Home Pages of researchers. Davis and Connolly further note that one of the challenges faced by the institutional repositories is that they need to encompass the many different normative disciplinary cultures represented by any one institution, a task that for now has proven difficult to accomplish. Xia and Sun (2007) have shown that even for the archival and preservation function that many institutional repositories are built for, the number of deposits is very low and many deposits in institutional repositories are not full texts. Further, most of the materials deposited in institutional repositories are done via proxy—deposited by library staff, research assistants, or departmental staff, rather than by the

researchers themselves. Numerous other research articles on institutional repositories address institutional repositories from the perspective of archiving and preservation of faculty's scholarly output rather than from the perspective of their use. Kennan (2008) has addressed the role of institutional open access repositories in scholars' knowledge production process in institutional settings, using ANT, by looking at numerous institutional actors. Similarly to the other studies however, she is interested in the contribution to the institutional repositories by the scholars and not in how the scholars use the materials found in them. Thus, considering that so far institutional repositories have not emerged as differentiating actors in researchers' knowledge production, and because the early attempts to recruit researchers that actively use institutional repositories to search, discover and access knowledge artifacts was not successful, only disciplinary repositories are considered for this study.

Given the distinct approaches to information seeking and use of digital resources between humanities scholars and scientists as noted by Palmer (2005), Palmer and Cragin (2008), Bates (1994), and Fry (2006), the participant selection process needs to identify researchers from the humanities and sciences disciplines.

The two possible scenarios are: a) humanities scholars accessing humanities OA disciplinary repository, and b) scientists accessing sciences OA disciplinary repository.

These scenarios will provide a rich set of data that will bring forth the interaction dynamics amongst the humanities scholars and scientists, and it will make possible to compare the dynamics between humanities scholars and scientists across the lived experiences, organizational and technological contexts as they interact with the open access repositories.

### 6.2.1 Methodological approach

To answer the research questions as set forth in this study, actors with their corresponding properties from the three contexts (lived experiences, technological, organizational) are taken into consideration for data collections and analysis.

Researchers' lived experiences context is extended and complemented by the technological and organizational context related to the access tools and open access repositories. While researchers' lived experiences will provide a social context that emphasizes on information practices and processes as they are experienced by the researchers, the technological and organizational contexts will provide insight into access tools' and OA repositories' technical capabilities (example: EPrints) and their organizational instantiations into specific repositories (example: arXiv).

As informed by the socio-technological co-constructionist approach these three contexts are linked and it is expected that the lived experiences context as expressed by the researchers will have references that point to the technological and the organizational capabilities related to the access tools and OA repositories. Reciprocally, the technological and the organizational contexts will have references to the lived experiences context in the form of expected and intended use of the technological and organizational layer elements as intended by the organizing structures.

As stated in the research questions, both the actors and their properties across the three contexts need to be investigated, as well as the links that connect these contexts. The data collection methods that will provide insight into researchers' lived experiences are guided by the phenomenological and grounded theory approaches. The data collection methods that will provide the insight into the technological and organization

capabilities of the access tools and OA repositories are guided by the actor-network theory.

The next two sub-sections describe and provide the rationale for the choice of data collection and data analysis approaches.

### 6.2.2 Data collection approach

This section addresses the congruency of data collection methods with the theoretical approach. The congruency between the research questions, interview questions, data collection and methodological approaches are also discussed.

Based on the phenomenological and grounded theory approach the study calls for understanding researchers' lived experiences as they interact with open access repositories. Researchers' lived experiences are situated within the larger socio-technological scholarly communication context and more specifically in the context that is enacted by the researchers, the access tools and the open access repositories. Understanding the access tools' and open access repositories' contexts requires investigation into their technological and organizational capabilities.

Given these distinct contexts, different data collection methods need to be applied. Researchers' most recent scholarly lived experiences as they access open access repositories will be collected through semi-structured interviews. Data about the technological capabilities and organizational structures of access tools and open access repositories will be collected through document analysis of documents that describe the open access tools and open access repositories. The data will contribute towards understanding three intertwined contexts: a) researchers' lived experiences and

perceptions, b) technological contexts related to the access tools and repositories, and c) the social and organizational contexts related to the access tools and repositories.

Baker, Wuest and Stern (1992) make a distinction between the data collection procedures appropriate for the phenomenological approach and the data collection procedures appropriate for the grounded theory approach, because phenomenology and grounded theory are based on distinct theoretical underpinnings. Baker et al. (1992) specifically note that the role of phenomenology is to understand and describe the lived experiences as presented by the participants (p. 1356). The role of grounded theory is to understand and describe the processes and sub processes that operate within a given context (p. 1357). Thus, the Research Questions (RQs) and the data collection procedures need to be carefully designed to correspond to the respective methodology. An attempt is made to build the RQs so that they can trigger the design of interview questions to solicit responses from the researchers to describe the essence of the phenomenon being studied as experienced by them, as well as to describe the processes that are enacted by the researchers as they interact with OA repositories.

Wimpenny and Gass (2000) take Baker et al.'s (1992) argument one step further and suggest congruency between research questions, interview questions, methodology and data collection. The interview questions as data collection instruments are designed to allow the participants of the study to describe their experience and also describe the processes. In the context of this study, the process can be also understood and learned by reading documentary evidence that pertains to the processes that are enacted or modified during researchers' interaction with the access tools and OA repositories. The goal of the document analysis is to provide a socio-technological perspective that may or may not

be easily visible by the researchers. The documents (such as About pages, Frequently Asked Questions (FAQs), and Help Guides) describe the processes that involve technology (access tools and repositories) and organizational structures. Together with the processes that emerge from the interviews with the researchers, it will provide the socio-technological and organizational contexts for researchers' lived experiences.

The data collection method for the interviews is informed by Barry (1995) who adopted the critical incident technique into information access stories that probes researchers' information access process by concentrating around researchers' article in progress as a pre-print or author's last published article. The critical incident technique around the article in progress (or a pre-print article), or last use of the repository for the purpose of producing a knowledge artifact, will trigger fresh memories of lived experiences and related processes due to proximity of the experiences and the focused interview questions: "Focusing on a particular paper provides a fairly direct means of access to memory, recovering information about specific information access methods which might otherwise not be remembered" (p. 131). In her study Barry investigates the implication of the electronic library and the use of information technology (IT) by researchers for information-seeking and research practices in academia. She accounts for researchers' information access methods and access to resources. From the perspective of context, relevant actors and process, Barry's study is congruent with the research design of this study.

### 6.2.3 Selection of participants

Considering the qualitative approach of this study with an emphasis on researchers' lived experiences, and the objective to understand and describe the common experiences

of a group of people that experience the same phenomenon, purposeful selection is used to identify the study participants. The initial set of participants is identified by examining the repositories for active participants. Some of the participants from the initial set may accept to participate in this study. The rest of the participants may be identified through the help of the initial set of participants, as well as administrator and maintainer of the repositories who might be familiar with active researchers. Thus, the participants will be researchers who are active users of OA repositories and use the OA repositories as resources to search, discover and access articles and other knowledge artifacts to be used for the production of scholarly work such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc. The process for participant identification is covered in section 6.3.3, The research participants.

#### 6.2.4 Analysis approach

Using grounded theory methods and through ANT's guided interpretations, the data emerging from the interviews and document analysis will reveal a wealth of actors and associations that need to be traced and brought together to enact contexts that can help understand the interaction between researchers, access tools, and repositories. Additionally, concepts, categories, themes and various properties and dimensions will emerge that will enable methodological and systematic description of the dynamics between the contexts.

The analysis will apply ANT's translation and inscription concepts as tools to describe how the technological characteristics and properties of access tools and repositories manifest themselves into the organizational aspects of access tools and

repositories, and further how both the technological and organizational aspects of access tools and repositories are manifested into the lived experiences enacted by the researchers.

As a result from the open and axial coding and linking concepts for relationships, the following models and contexts guide the analysis to understanding the implication of the open access resources on researchers' individual and disciplinary knowledge production contexts:

1. *Lived experiences and perceptions contexts (individual and group level)*

The lived experiences context will mostly emerge from the interviews in terms of processes, relevant actors and linking structures. The interviews will also provide links into the technological and organizational context of access tools and repositories by identifying the relationships that link actors from the lived experiences context with the technological and organizational context.

2. *Technological context*

The technological context for the access tools and the repositories will mostly emerge from the document analysis in terms of process, relevant actors and linking structures. The document analysis will also provide links into the organizational context for the access tools and repositories and also provide links into researchers' lived experience context.

3. *Organizational context*

The organizational context for the access tools and the repositories will mostly emerge from the document analysis in terms of process, relevant actors and linking structures. The document analysis will also provide links



into the technological context topologies for the access tools and repositories and also provide links into researchers' lived experience topology.

4. *The knowledge production context - spanning across the three contexts (individual and group level)*

These contexts will be developed at individual level as well as disciplinary level and will reveal the dynamics between the three contexts.

Each context will reveal a set of actors, their relationships and their relevant properties. Within each set of topologies, patterns of interactions will emerge.

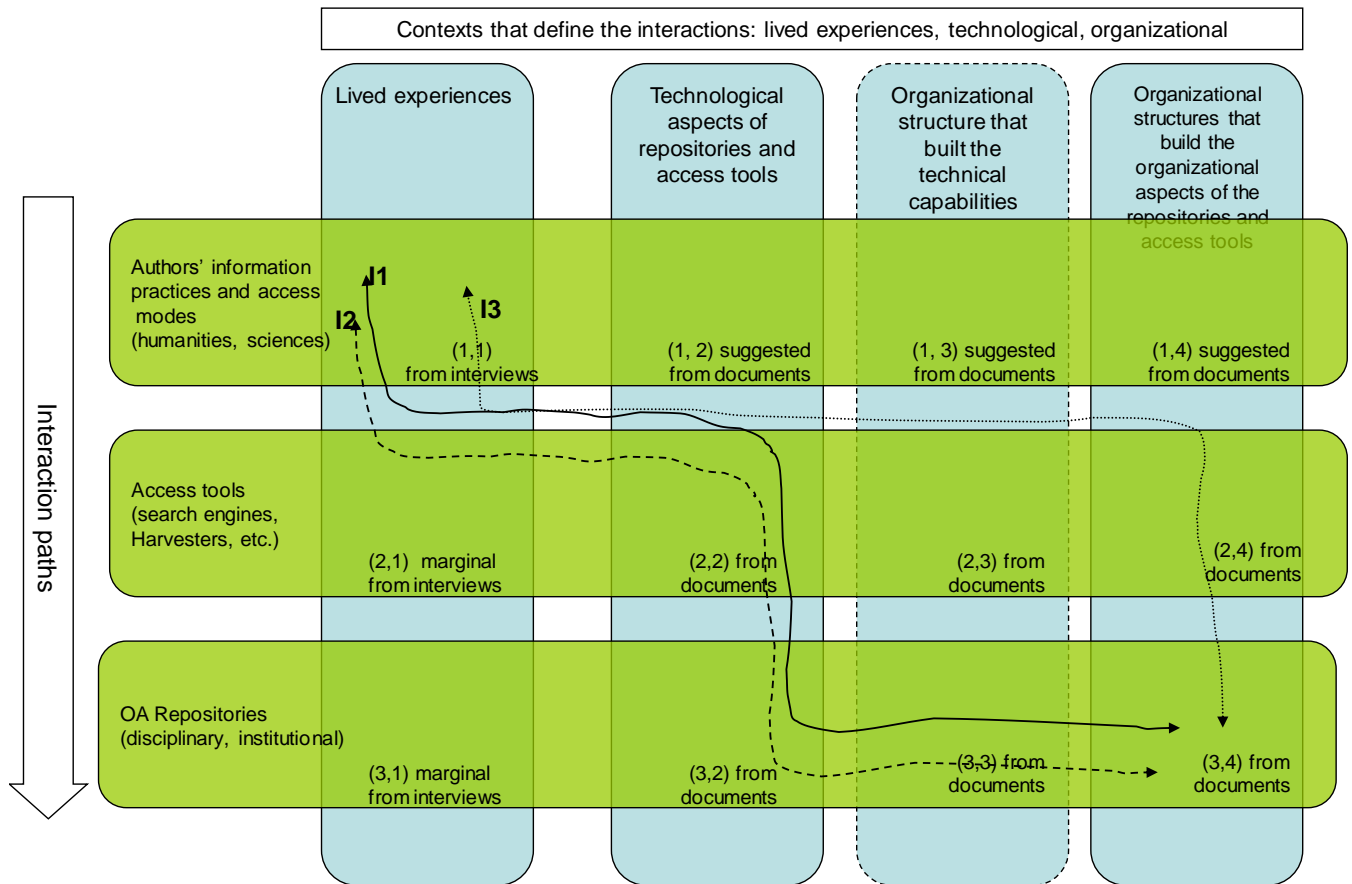
Figure 4 describes the paths of possible interactions around which the interview questions are constructed. The interactions enact and trigger dynamics between the horizontal layers that represent the researchers, access tool, and the repositories; and the vertical layers that represent the three contexts: lived experiences, technological and organizational. The intersections between the vertical and horizontal layers denote data collection areas that will contribute towards the understanding of the relationships between the three contexts. The intersections are denoted by the (x, y) label and the data collection methods is noted. For example, the data in intersection (1, 1) is collected through the interviews with the researchers.

To further contextualize the interactions, the following describes the data collection method that provides the data for the intersections between the main actors (researchers, access tools, repositories) and the three contexts (lived experience, technological, organizational):

- (1,1) - Primary source for the lived experience context. Data will emerge from the interviews with the researchers.

- (2,2), (2,3), (2,4) - Primary source for the technological and organizational aspects of the access tools. Data will emerge from the document analysis.
- (3,2), (3,3), (3,4) – Primary source for the technological and organizational aspects of the OA repositories. Data will emerge from the document analysis.
- (2,1), and (3,1) – Marginal description of access tools and repositories as perceived by the researchers. Data will emerge from the interviews.
- (1,2), (1,3), (1,4) – Marginal and perhaps suggested use of access tool and repositories by the researchers. Data will emerge from document analysis.

Figure 4: Interactions as triggers for scholar's lived experiences and perceptions



- Each intersection (x,y) between the horizontal and vertical layers contains many actors

- I1, I2, I3 – sample interaction paths. Each interaction is instantiated by a set of actors .

I1, I2, and I3 in Figure 4 are example interactions enacted by the researchers, intended to show the possible dynamics that may emerge between the lived experiences and perceptions layers, the technological layers and the organizational layers. As researchers interact with open access repositories, actors and properties from the technological context and the organizational contexts of the access tools and repositories are put into motion. For example, interaction I1 denotes a researcher's lived experience that is predominantly impacted and determined by the technological properties of the open access repositories. Interaction I2 denotes a researcher's lived experience that is predominantly impacted and determined by the technological properties of the open access repositories and the access tools. And, I3 denotes a researcher's lived experience that is predominantly impacted and determined by the organizational structures of the open access repositories.

### 6.3 *The Research Process*

Guided by the theoretical approach and the methodological considerations as discussed in the previous sections, the rest of this sub-section presents the identification and selection process of OA repositories and participants, and also the design of data collection instruments.

#### 6.3.1 Data collection

Based on the understanding from the existing literature that researchers as authors, access tools and self-archived OA repositories are relevant actors situated in a socio-technological context, the data collection approaches need to account for both the understanding of researchers' lived experiences and the socio-technological processes

being enacted. Thus, two data collection approaches are used: interviews and document analysis. Researchers are interviewed with the intent to solicit information about their lived experiences and perceptions around how they discover and access the knowledge artifacts needed for their knowledge production process. The interview guide will also probe their everyday information practices as they go about finding knowledge artifacts, how they approach various access tools, etc. The document analysis of access tools related documents would provide the understanding about their metadata definitions, user-facing interfaces, system facing interfaces, etc. This will reveal a set of central actors and links from the perspective of the access tools. Similarly, the document analysis of repository related documents would reveal actors and links that enable the interaction with users as well as the interaction with other systems.

Before the participants can be recruited, the repositories needed to be identified first.

### 6.3.2 Identifying OA repositories

One of the tasks for this study is to identify a combination of researchers and OA repositories that will provide authentic lived experiences of researchers that are active users of OA repositories. An active user is a researcher who uses knowledge artifacts from OA repositories to produce knowledge artifact of his/her own such as scholarly articles. In this section, active sciences and humanities OA disciplinary repositories are identified. After identifying and selecting the repositories, the approach for the identification and selection of participants is presented.

The Registry of Open Access Repositories (ROAR) at <http://roar.eprints.org/> is a good starting point as it provides a systematic way to identify OA repositories based on specific selection criteria. The main four criteria were: geographic (by country), types of

software systems, content type (includes categories of repositories such as Research Cross-institutional, and Research Institutional or Department), and sorting (name, number of records, time of last activity, and activity load). The Research Cross-Institutional type label used by ROAR is how ROAR labels OA disciplinary repositories.

As of January 27<sup>th</sup>, 2008, ROAR had a total of 989 repositories, out of which 225 are in the United States. The identification of repositories was limited to those located in the United States to avoid language barriers with the participants during the interviews as well as the ability to manage cost for conducting the study. Further, out of the 218 repositories in the US, 16 were research cross-institutional (i.e., disciplinary) and 127 institutional or departmental.

To identify the repositories, ROAR was used to list the disciplinary repositories in the United States, and then the list was sorted in a descending order of number of records to identify those repositories with the highest number of records as potential indication for more active repositories. The resultant list, shown in Appendix B (see Table B1), was scrutinized to identify the repositories that contained peer-reviewed articles and to identify the disciplines.

In addition to ROAR, Appendix B also lists OA repositories based on The Directory of Open Access Repositories (OpenDOAR) project at <http://www.opendoar.org>. The same selection criteria are used as in ROAR listing in Table B1, with the addition of the type of content listing criteria that was available in OpenDOAR. This made it easier to determine whether a repository contains articles, theses or other types of material. Table B2 shows the list of disciplinary (cross-institutional) repositories based on OpenDOAR.

Table 2: Selected disciplinary repositories

<b>Discipline</b>	<b>Repository name</b>	<b>Brief description from the repository portal</b>	<b>Location</b>
Sciences	Arxiv.org	arXiv is an e-print service in the fields of physics, astronomy, astrophysics, mathematics, non-linear science, computer science, and quantitative biology.	Cornell University Cornell, NY  <a href="http://www.arxiv.org/">http://www.arxiv.org/</a>
Humanities	PhilSci Archive	Philosophy of Science Archive	University of Pittsburgh Pittsburgh, PA  <a href="http://philsci-archive.pitt.edu">http://philsci-archive.pitt.edu</a>

The identification process as described above resulted in the following selection, shown in Table 2, of disciplinary repositories for sciences and humanities disciplines:

1. *Sciences*

arXiv (<http://www.arXiv.org>) is an obvious selection considering that it is one of the earliest disciplinary archives, has a large number of articles written about it, is very active, and a number of documents describing its capabilities are available for document analysis. arXiv is listed in both ROAR and OpenDOAR.

2. *Humanities*

Identifying disciplinary OA repository in humanities proved to be more challenging. It is possible that disciplinary OA repositories in the humanities are in the early stages of development due to the nature of how scholarly knowledge is distributed: most of humanities scholarship is based on books and to a lesser extent on articles. Nevertheless, both ROAR and OpenDOAR lists the Philosophy of Science Archive (PhilSci, <http://philsci-archive.pitt.edu>) OA repository as the only active humanities disciplinary repository that contains

articles (pre-prints or post-prints). It has a substantial number of pre-print and post-print articles related to different aspects of philosophy of science. It is not listed in Table B2 since it is 36<sup>th</sup> in descending order based on the number of records it contains. However, ROAR lists PhilSci with 1530 number of records.

Based on the preceding findings and identification of repositories, these repositories are briefly described so that they can inform and contextualize the participant selection process. The following are some of the criteria and delineations used in the description: a) brief history, b) scope, purpose, and goals, c) submission policies, d) types of materials deposited therein, e) organizing entities, f) the software that runs the repository, and g) the interoperability and metadata standards.

#### 6.3.2.1 arXiv

arXiv.org (formerly known as xxx.lanl.gov) is the longest running disciplinary self-archive repository. It has established itself as a critical actor in the knowledge production in scholarly disciplines such as physics, high energy physics, astronomy, astrophysics, mathematics, computer science, nonlinear science, quantitative biology and statistics (About arXiv). It was started as an experiment by Paul Ginsparg in 1992 running on a single computer in his office. Now it is run by the Cornell University Library with guidance from the arXiv Advisory Board (About arXiv). Over the years, through a systematic approach with the goal to make pre-prints of scholarly work accessible to the relevant scholarly community, arXiv has reached a status that is perhaps unmatched by other OA repositories in terms of the amount of pre-prints it contains and its established role as a source for scholarly knowledge artifacts for the specific disciplines it covers. Most of the artifacts deposited in arXiv are pre-prints of

articles, often multiple versions, which have been already accepted for publication in peer-reviewed journals or are of high value to the specific scholarly community (About arXiv).

Some of the stated goals for arXiv are (from arXiv Primer):

- a) Supplement the traditional scholarly publication systems by providing immediate dissemination and open access to research articles; i.e., arXiv is not an archive of unpublished materials
- b) A moderation process is established to ensure that submissions are of value to the arXiv communities
- c) Preserve the scholarly records, i.e., submissions are not removed after they are deposited
- d) A commitment to provide *persistent access* and *perpetual availability* to all announced submissions
- e) “...Material submitted to arXiv is expected to be of interest, relevance, and value to those disciplines. arXiv reserves the right to reject or reclassify any submission. Submissions are reviewed by expert moderators to verify that they are topical and refereeable scientific contributions that follow accepted standards of scholarly communication.”

arXiv has achieved the role of quality and value by instantiating a moderation process by which established and published researchers with higher number of submissions into the repository review the submissions. But, for a scholar to be able to submit to arXiv, the scholar needs to be endorsed by an active member (an established member who has deposited a number of articles to arXiv for few years) related to the



specific arXiv community. As the arXiv Primer states, the purpose of the moderation process is to provide interested researchers with relevant and legitimate research as free resources: “This process helps restrict arXiv submissions to relevant and legitimate research contributions without adding to the administrative cost of arXiv, and thus it is an essential contribution to both the legitimacy and the sustainability of arXiv as a free resource.”

arXiv is implemented as a “home built” system of software components. Based on the interface and portal it seems like a variation of the EPrints software which is more common software for hosting OA repositories. The articles in arXiv can be accessed directly through its own interface at <http://www.arxiv.org>, or indirectly through generic search engines such as Google, scholarly specialized engines such as Google Search and Scirus, as well as meta-search engines. arXiv supports OAI Dublin Core (DC) metadata format and the OAI-PMH interoperability protocol for metadata exchange.

#### 6.3.2.2 PhilSci

The PhilSci OA disciplinary repository was established in 2000 by Rob Clifton, John Earman and John Norton, professors in the Department of Philosophy and History and Philosophy of Science at the University of Pittsburgh. Officially opened for submissions in 2001, PhilSci is supported, endorsed, and run by the Center for Philosophy of Science, the office of the Dean of the Faculty of Arts and Science, the Philosophy of Science Association and the University Library System of the University of Pittsburgh (About PhilSci).

Some of the stated goals of PhilSci are (from PhilSci Policy):

- a) To be sustained as a free service to philosophers of science

- b) Preserve and foster rapid exchange of new work in the philosophy of science
- c) Act as a gatekeeper to ensure that only materials that are of interest to the professional philosophers of science are deposited therein. The style and topics are guided by what is potentially publishable in the Philosophy of Science journal
- d) Only scholarly papers that have achieved stable form are posted therein
- e) PhilSci has a policy that, as per author's request, files and deposits in PhilSci can be removed

Unlike arXiv, any philosopher of science can register with PhilSci and start posting their articles immediately; there is no need for newcomers to PhilSci to be endorsed by anybody. The articles are reviewed by the moderators of PhilSci for topical and thematic relevancy to philosophy of science.

PhilSci is also compliant with OAI-PMH. As such, articles in PhilSci can be accessed through its own portal at <http://philsci-archive.pitt.edu> or through other general or specialized search engines, or through meta-search engines that harvest the OAI DC compliant metadata records from PhilSci.

The difference between archives' own search and portal access and other search engines (Google Scholar, meta-search engines) is that through the portal you can search the entire content of the submissions, while meta-search engines might be limited to searching the metadata only.

PhilSci, like arXiv, has a policy in place that allows the harvesting of metadata by default but do not allow harvesting of the full text of articles without prior agreement. PhilSci uses the EPrints software to run its repository.

### 6.3.3 The research participants

As stated earlier in this section, the initial criterion for identifying the researchers is that they actively use OA repositories (one or many) in their knowledge production. To identify the active researchers—users of and contributors to OA repositories, each repository was scanned separately for a set of participating researchers at the high end of number of article submissions (i.e., using purposeful selection). However, there are some challenges with this approach. It is not clear whether high submitting researchers also use the repository to discover and access artifacts for their knowledge production. Further, the proportion of articles (per-prints and post-prints) and theses in the identified disciplinary OA repositories from ROAR and OpenDOAR (see Table B1 and Table B2) are not easily ascertainable. On the other side, the disciplinary repositories mostly contain articles and only few contain theses.

In order to understand a little more about what types of submissions the selected repositories contain and who deposits in them (researchers or somebody on their behalf), the metadata records (DC format through OAI-PMH interface) from these repositories were downloaded. *Table 3* shows simple descriptive statistics with respect to the content deposited.

Table 3: Brief statistics for the selected repositories

Repository	Repository type	OAI interface URL	Number of records	Type of records
arXiv	Sciences	<a href="http://export.arxiv.org/oai2">http://export.arxiv.org/oai2</a>	- 460118 total records	- pre-prints - post-prints - other
PhilSci	Humanities	<a href="http://philsci-archive.pitt.edu/perl/oai2">http://philsci-archive.pitt.edu/perl/oai2</a>	- 1842 total records - 286 deleted - 1556 active total	- 270 conferences papers - 1095 preprints - 191 other

With respect to arXiv, Brown (2001) notes that most records in arXiv are pre-prints and post-prints, self-submission is encouraged, and that there is no provision for proxy submission by others on behalf of the researchers (arXiv materials are submitted by the researchers themselves). With respect to PhilSci, a large number of random articles were browsed via PhilSci's portal and all of them were deposited by the researchers themselves. The browsing via the PhilSci portal was a substitute for the lack of differentiation of researchers vs. depositors in the PhilSci OAI DC metadata records.

Given the large number of researchers that use arXiv and PhilSci, the next step in the process was to identify specific researchers. The large research universities on the east coast of United States became the starting points for the identification of active researchers in arXiv and PhilSci. This was intended to help with the logistics for conducting this study by reducing cost and at the same time to provide a pool of researchers that are easier to access due to proximity of location.

#### 6.3.3.1 Identification and selection of arXiv scientists

Considering that the identification of researchers who use arXiv to access materials needed for their knowledge production is not directly possible, the identification of researchers who contribute to arXiv is used as a proxy. The identification of arXiv participants proceeded in two stages. At the beginning of the first stage, high contributing authors from four large research universities on the eastern coast of United States were targeted. This produced four (4) scientists with high contribution of materials in arXiv. All of the materials were deposited by the scientists themselves. At this point, these scientists are treated as potential participants, and they can also help with identifying other potential participants that are active users of arXiv. The RFP

(Request for Participation, Appendix C) was sent directly to these four potential participants, to their publicly available e-mail address. Only two of the participants responded agreeing to participate in the study.

Then, the potential list of participants was expanded to include potential participants from other research institutions throughout the United States. A total of sixteen (16) potential participants were identified. The RFP was sent directly to their publicly available e-mail addresses. As a second stage, a discussion forum community related to physics and related disciplines was identified as potential forum where arXiv users may participate. The RFP was sent to PAMnet@listserv.nd.edu (available at <http://listserv.nd.edu/cgi-bin/wa?A0=PAMNET>), with a user base of about 632 subscribers.

These two approaches resulted in six (6) participants accepting to participate in this study: three (3) from large research universities in the United States, one (1) from a medium research university in the United States and two (2) from large observatories in the United States. Interestingly, all participants were either astronomers or astrophysicists. The participant from the medium research university in the United States learned about the RFP via the science librarian who participates in PAMnet. The two participants from the large observatories in the United States, of whom one is in the last stages of her Ph.D. studies, received the RFP through their librarian who also participates in PAMnet.

#### 6.3.3.2 Identification and selection of PhilSci scholars

The identification and selection of PhilSci participants was initiated by downloading the OAI DC metadata records from the PhilSci repository via its publicly accessible

OAI-PMH interface. The downloaded records were then sorted using researchers' names as the sorting key. The list of researchers was searched for instances of submissions in the repository directly from its own portal. A number of active scholars (based on higher number of submissions) were identified from large research universities across the United States. As with arXiv, these PhilSci identified scholars are treated as potential participants that can also help with the identification of other participants. In the first stage, based on the proxy criteria that participants are active high contributors to PhilSci, thirty three (33) potential participants were identified. The RFP was sent directly to their publicly available e-mail addresses. The initial response was slow. Thus, as a second stage, the administrators of PhilSci (both academic and technical) were contacted and asked for help and advice about how to best identify active users of PhilSci for this study. The response from the PhilSci administrator was not very helpful in the identification of participants. So, in addition to directly sending the RFP to the potential participants, a scholarly discussion community related to the subject of History of the Philosophy of Science was identified as potential forum where PhilSci users may participate. The RFP was sent to HOPOS-L@listserv.vt.edu (available at <http://listserv.vt.edu/archives/hopos-l.html> ), with a user base of about 1001 subscribers.

After few attempts of resending the RFP and direct telephone communication at their workplace publicly available phone numbers, five (5) participants accepted to participate in the study. Four (4) of the participants were identified by direct communication (via e-mail and telephone). One (1) participant, a graduate student, was referred to the RFP by another graduate student who was identified via his personal blog due to his interest in

philosophy of science and his expressed interest around the value that PhilSci brings to philosophers of science.

#### 6.3.3.3 Summary of the participant identification process

The participant selection process, a combination of the snowball and purposeful selection methods, resulted in 11 participants—6 arXiv scientists and 5 PhilSci scholars, with 9 participants across the United States, and 1 each in Canada and the United Kingdom. This number is adequate for a qualitative study that emphasizes participants' lived experiences and the objective to understand and describe the common experiences of a group of people that experience the same phenomenon. This is in congruence with the study design as the goal is to identify participants that experience the same phenomenon, i.e., use of OA repositories. In this case, all of the participants that responded and accepted to participate self-identified themselves as active users of materials from OA repositories. The self-identification step for the potential participants was embedded in the RFP. The following text is included under the Participant Characteristics section of the RFP: “The participants will be scholars who are active users of OA repositories and use the OA repositories as resources to find, discover and access articles and other knowledge artifacts to be used in their knowledge production.”

The characteristics of the researchers that agreed to participate in this study are listed in Table 4.

Table 4: arXiv and PhilSci participant characteristics

Participant code	Institution type and department affiliation	OA archive	Date Ph.D. received	Academic rank or position	Gender
A1	Large research university in the United States <i>Department of Physics and Astronomy</i>	arXiv	1984	Professor	Male
A3	Large research university in the United States <i>Department of Physics and Astronomy</i>	arXiv	1977	Professor	Male
A17	Large research university in the United States <i>Department of Physics and Astronomy</i>	arXiv	1998	Assistant Professor	Male
A18	Large observatory in the United States <i>Office of Operations</i>	arXiv	Graduate student at the time of the interview	Assistant Director	Female
A19	Medium research university in the United States <i>Department of Physics</i>	arXiv	2004	Assistant Professor	Female
A20	Large observatory in the United States	arXiv	1996	Tenured Astronomer	Female
Ph10	Large research university in the United States <i>Department of History and Philosophy of Science</i>	PhilSci	Graduate student	NA	Female
Ph24	Large research university in the United States <i>Department of Philosophy</i>	PhilSci	2007	Assistant Professor	Male
Ph26	Major research university in the United Kingdom <i>Department of Philosophy Logic and Scientific Method</i>	PhilSci	2003	Senior Lecturer	Male
Ph29	Large research university in Canada <i>Department of Philosophy</i>	PhilSci	2006	Assistant Professor	Male
Ph32	Medium master's level university in the United States <i>Department of Philosophy</i>	PhilSci	1994	Professor	Male

Note: The non-sequential codes for the participants signify the order number by which they were identified as potential participants for this study. The non-sequential numbering has been kept throughout the dissertation for practical purposes since the codes were used for coding and analysis before the full recruitment was complete. For the institutions based in the United States, the institution type was determined using the Carnegie Classification of Institutions of Higher Education. For the institutions based outside of the US, the institution type was ascertained by reading the “About” pages of the respective institution and its relation to the other educational institutions in their respective country.

All six arXiv participants are astronomers or astrophysicists who mostly interact with the physics subject area of arXiv, most specifically with the astro-ph category that contains scholarly materials related to: cosmology and extragalactic astrophysics, earth



and planetary astrophysics, galaxy astrophysics, high-energy astrophysical phenomena, instrumentation and methods for astrophysics, and solar and stellar astrophysics. Throughout this study, references to the arXiv content and its scope refer specifically to the astro-ph category only. However, the search, browsing, linking and other features and capabilities are the same across the different subject areas represented in arXiv.

#### 6.3.4 The interview process

This section provides the rationale for designing the instruments for data collection, especially the construction of the pre-interview questionnaire and the interview guide. It also describes the process for conducting the interviews.

##### 6.3.4.1 Pre-interview questionnaire

The primary goal of the pre-interview questionnaire was to establish the background about researchers' work, their understanding of the role of OA in their field of study and to identify their recently produced knowledge artifacts or work in progress around which some of the interview questions are related.

The pre-interview questionnaire was sent to the researchers together with the consent form (shown in Appendix D) and the interview questions (shown in Appendix F). The pre-interview questionnaire is included in Appendix E.

##### 6.3.4.2 Interview questions

The following guidelines and perspectives were carefully considered and informed the design of the interview questions to be used for the semi-structured interviews with the researchers:

- As informed by the phenomenological approach, the questions should trigger responses from the participants to describe their lived experience.
- As informed by the grounded theory approach, the questions should trigger responses from the participants to describe the processes within which or by which the lived experiences are enacted.
- As informed by the co-constructionist approach, the questions should trigger socio-technological inclusiveness both for the lived experiences and for the processes.
- As informed by the actor-network theory, the questions should solicit responses that describe the relationships and links between various socio-technological actors and processes.
- Based on Barry's (1995) information access story approach, and the critical incident technique (Flanagan, 1954, p. 337), the interview questions will be focused around researchers' recently produced knowledge artifacts or works in progress, with the aim to understand the process and the steps that researchers engage in discovering and accessing knowledge artifacts in open access repositories.

The interview questions are included in Appendix F. These interview questions were the starting point for interviewing each participant. However, researchers' responses to the pre-interview questions were used to modify and clarify the interview guide (Appendix G), especially by identifying additional prompts for the interview questions for each individual researcher.

The sequence of the interview questions as listed in the interview guide is meant to trigger responses to lived experiences by the participants regarding their use of OA repositories and access tools, from broader to a more specific perspective. In order to contextualize participants' lived experiences within the broader context, Q1 and Q2 are meant to provide broader information about participants' involvement with OA repositories in general and also to probe their perception about the role of OA in scholarly communication. Q3, Q4, and Q5 are meant to trigger specific memories around the OA repositories the participants use regularly, their recent knowledge artifacts and the role OA repositories play in the different stages of their knowledge production. Q5 and Q9 should trigger responses that would compare the role of resources found in OA repositories with those found elsewhere, as used in different stages of knowledge production. Q6 is meant to trigger important uses of OA articles by the participants in addition to referencing them in the article they are writing. Q7 and Q8 are expected to trigger responses around the ways the participants search, discover and access articles in OA repositories. Q10 should reveal any other OA repositories used by the participants in addition to the specific repository that they have been identified for this study. Q11 is meant to trigger experiences around specific challenges when using OA repositories. The goal is that such strong lived experiences will reveal important processes that go beyond surface description of experiences. Q12 aims to reveal an information access process enacted by the participants when they know a certain reference or a quote exists but they cannot remember the reference exactly. How do participants go about finding and discovering such references? What is the role of OA repositories and access tools in this case? Q13 aims to discover challenges with

participant's information access process by probing for any need for improvement with access tools and OA repositories. Finally, Q14 is a catchall question that provides the participant with an opportunity to talk about any issues they perceive important but were not addressed by the interview questions. The interview guide (Appendix G) contains a number of prompts for each question that helps to guide the interview by triggering descriptions of relevant and sometimes specific experiences.

There is stronger emphasis on researchers' lived experiences and related processes in the interview questions, and less emphasis on the technological and organizational characteristics of the access tools and OA repositories. However, based on the interview questions, it is expected that the participants will provide some description about the technological and organizational aspects of access tools and repositories as they describe their interactions. The organizational and technological aspects that are mentioned by the researchers are used as guidance about where to look for specific access tools and repository capabilities that subsequently will be collected through document analysis.

#### 6.3.4.3 Conducting the interviews

The interview questions were shared with the participant beforehand. The interviews were conducted using the interview guide. The interviews ranged from 40-75 minutes in length and were conducted over a period of five months. The interviews were taped with a digital handheld recorder and were subsequently transcribed by a professional transcription service. Participants A1, A3, A17, A19 and Ph32 were interviewed in person at their office, with the exception of Ph32 who was interviewed in a conference room. The rest of the participants were interviewed over the phone.

#### 6.3.4.4 Pilot study

The pilot study was limited to one participant to provide feedback around the flow of the interview process, including the pre-interview questionnaire, conducting the interview, test the length of the interview, as well as to assess whether the interview questions triggered relevant lived experiences as well as data related to the technological and organizational aspects of open access repositories and access tools. The selection criterion was that the participant be a scholar who is an early adopter of OA repository use.

After transcribing the interview, the responses to each question were compared with the expected outcomes in terms of lived experiences, processes and relationships between lived experiences and links to the technological and organizational contexts. By and large, the interview provided valuable learning experiences and the interview questions triggered the desired responses to provide valuable insight into the lived experiences and links into the technological and organizational contexts. The interview guide (Appendix G) was modified to include specific prompts based on the learning from the pilot study. The pilot study also helped me become more comfortable with the interview process, and it helped me improve my interviewing skills.

#### 6.3.5 Documentary evidence: identification and analysis procedures

The following guidelines and perspectives were carefully considered and have informed the document analysis process, mainly informed by the co-constructionist approach. The goal of these guidelines was to provide a direction towards the identification of documents that describe the socio-technological context and the

processes within which researchers' interactions with access tools and OA repositories are enacted. The guidelines also act as the guiding boundary for document collection.

- a) Documents that describe the technical constructs related to the access tools and OA repositories and their capabilities, as articulated and described by the organizing structures:

- a. Software with which these repositories and access tools are built.

*Examples:* EPrints and Google Scholar.

- b. The metadata standards and interfaces supported by the access tools and OA repositories.

*Examples:* OAI DC metadata, Extensible Markup Language (XML) schema, OAI-PMH protocol and implementation guidelines, etc.

- b) Documents that describe the organizational structures that setup and maintain arXiv and PhilSci as social structures, i.e., the non-technical aspects of the OA repositories, such the selection process of what is included and excluded in the repositories, funding for ongoing maintenance, thematic structures, policies, etc.

*Example:* Cornell University Library consulted by the arXiv Advisory Board members decides on policy issues for arXiv.

The collected documents for analysis are primary sources such as design documents, feature description documents, About and FAQ pages, etc. Most of the documents in Table 5 were identified based on the literature review. Others, such as those describing Astrophysics Data System (ADS), were identified based on the learning from the interviews with the participants.

The selection and identification process for document collection was as follows:

- a) Based on the repository selection, the software and the associated metadata and interface standards for arXiv and PhilSci are identified. Probable locations for these sets of documents are the portals of the OA repositories, the portals of the software that is used to build the repositories, and the websites that describe the metadata standards and the interfaces. This step will provide documents that:
  - a. Describe the technical aspects of the software used to build the repositories and their capabilities.
  - b. Describe the organizational structures that build these repositories.
  - c. Describe the technical aspects of the metadata and the interfaces.
  - d. Describe the organizational structures that build the software, the metadata and the interfaces.
- b) Based on the technical understanding of the repositories, the metadata and interfaces, in conjunction with the findings from the interviews where the researchers identify the access tools they use, a number of access tools will be identified. This step will provide documents that:
  - a. Describe the technological aspects of the access tools and their interoperability with metadata and interface standards.
  - b. Describe the organizational structures that build the access tool.

The selection process outlined above resulted in the following documents shown in Table 5.

Table 5: Document analysis sources

<b>Doc name and ref source</b>	<b>Repository / access tool / software</b>	<b>Name of repository, access tool, or software</b>	<b>Technological or organizational, layer</b>	<b>Genre / document type</b>	<b>Target audience</b>
<a href="#">About arXiv</a> retrieved on 1/29/2009	Repository	arXiv	Organizational	About	Discoverers and contributors of OA
<a href="#">arXiv Primer</a> retrieved on 1/29/2009	Repository	arXiv	Organizational	A guide	Discoverers and contributors of OA
<a href="#">arXiv FAQ</a> retrieved on 1/29/2009	Repository	arXiv	Organizational Technological	FAQ	Discoverers and contributors of OA
<a href="#">About PhilSci</a> retrieved on 1/29/2009	Repository	PhilSci	Organizational	About	Discoverers and contributors of OA
<a href="#">PhilSci Policy</a> retrieved on 1/29/2009	Repository	PhilSci	Organizational	Policy doc	Discoverers and contributors of OA
<a href="#">PhilSci FAQ</a> retrieved on 1/29/2009	Repository	PhilSci	Organizational Technological	FAQ	Discoverers and contributors of OA
<a href="#">PhilSci Help</a> retrieved on 1/29/2009	Repository	PhilSci	Organizational Technological	Help and User Guide	Discoverers and contributors of OA
<a href="#">EPrints 2.2 Documentation</a>	Repository Software	EPrints	Technological	Administrator's Manual	Scholarly Communities
<a href="#">NASA/SAO/ADS FAQ</a> retrieved on 1/29/2009	Access tool, aggregated catalog of resources	ADS	Technological Organizational	FAQ	Discoverers and contributors of OA
<a href="#">A brief description of the ADS, and a disclaimer – User Guide</a> retrieved on 1/29/2009	Access tool, aggregated catalog of resources	ADS	Organizational	User Guide Brief description	Discoverers and contributors of OA
<a href="#">SAO/NASA ADS Help Pages</a> Retrieved on 9/12/2009	Access tool, aggregated catalog of resources	ADS	Technological	Help Pages	Discoverers and contributors of OA
<a href="#">About Google Scholar</a>	Access tool	Google Scholar	Technological	About	Discoverers and contributors of OA
<a href="#">History and development of OAI-PMH</a> retrieved on 1/29/2009	Protocol, interface	OAI-PMH	Technological Organizational	Tutorial	Data providers, service providers, archives, policy makers, software designers, interface designers, etc.



<a href="#">OAI FAQ</a> retrieved on 1/29/2009	Standards	OAI	Technological Organizational	FAQ	Data providers, service providers, archives, policy makers, software designers, interface designers, etc.
<a href="#">About OAI</a> retrieved on 1/29/2009	Standards	OAI	Technological Organizational	About	Data providers, service providers, archives, policy makers, software designers, interface designers, etc.
<a href="#">OAI and OIA-PMH Overview</a> retrieved on 1/29/2009	Standards, Protocol, Interface	OAI OAI-PMH	Technological Organizational	Tutorial	Data providers, service providers, archives, policy makers, software designers, interface designers, etc.

Note: Some of these sources were identified based on the literature review. Others, such as documents related to ADS were identified based on the learning from the interviews with the scientists.

The documents shown in Table 5 are primary documentary evidence only. These are documents that are articulated and written by the organizing and management structures that build, maintain and construct either the technological layers or the organizational layers of the open access repositories and the related access tools. The table does not contain secondary documentary evidence in the form of research articles.

Out of the sixteen (16) documents, four (4) are About pages, four (4) are FAQs pages, seven (7) are Help pages, guides and tutorials, and one (1) policy document. It is important to note that the type of content these documents provide is not clearly delineated across the documents, and that the title is not necessarily a clear indicator of the content type.

The documentary evidence was analyzed by close examination of the text from which the quotes related to the organizational and technological layers of the open access repositories, the access tools and the relevant standards and protocols were

extracted. The quotes were then analyzed to reveal the relevant actors and their properties and to understand the technical level features and capabilities as well the organizational level intended goals and values—both of which were compared between those used by the astronomers and those used by philosophers of science.

#### 6.3.6 Software for qualitative analysis

NVivo is used for the coding and analysis of the transcribed interviews. Needless to say, the identification of categories (concepts, actors, themes, and relationships), their properties and dimensions, and the processes from the data interviews is not a mechanical process using this software. The software was used iteratively by the understandings that emerge from the data itself throughout the data analysis and interpretation stages, especially with respect to the granularity of coding. The constant comparison method as an iterative process is one of the ways to ensure that categories and processes are coded with a balanced approach, not too detailed but also not too abstract, in order to explain the research questions at the appropriate level. In relation to the grounded theory, NVivo is used iteratively for open coding, axial coding and selective coding.

NVivo was not used for the analysis and interpretation of the organizational and the technological layers from the documentary evidence. Instead, due to the comparatively lower volume of content (i.e., documentary evidence) in comparison with the interviews with the researchers, these procedures were carried out manually, in multiple iterations, by tabulating the features and functional capabilities (technological level) and the intended roles and values (organizational level), respectively for the two groups of

participants. The analysis and interpretation of the documentary evidence is presented in Chapter 11.

#### *6.4 Interviews: coding and analysis procedures*

Data coding and data analysis is an iterative process that happens concurrently when using the grounded theory approach. There is an emphasis on the constant-comparison task, an iterative process that helps refine, clarify, link, and balance the identification of categories, concepts, themes, and actors at the appropriate level of understanding (concrete and conceptual level).

The interview coding and analysis proceeds as follows in two stages: a) writing memos, taking notes, and high-level summarization of the interviews, and b) open, axial, and selective coding of the interviews.

##### *6.4.1 Memos and notes*

Stage one started with the notes and memos written right after the interviews were conducted. Once the transcriptions of the interviews were completed, each interview was then summarized in a three-page document. Further, the combination of memos, notes and the three page summary document were further analyzed to produce: a) a one page concise summary for each scholar that describes their information processes, and provides a summary of their perceived value and role that OA repositories and access tools play in researchers' knowledge production, and b) for each scholar a summary of notes, thoughts, emerging ideas, and some preliminary analysis and interpretations were written reflecting investigator's interplay with the process up to that point.

Guided by Strauss and Corbin (1998, p. 217-23) and Miles and Huberman (1994, p. 72-75), the goal of these memos, summaries and reflections was to have an easily visible view of each scholar and to keep track of categories, concepts, themes, relationships, thoughts and ideas that emerge during the research process—all this with the intention that contextual reflections that are captured will be helpful in elucidating interpretations and help with the grounded theory methods.

Concurrently with this stage (i.e., memos and notes), the process continues with the methodological approach of data coding and analysis.

#### 6.4.2 Interview coding procedures

In line with the grounded theory approach, data coding was performed using open coding, axial coding and selective coding methods. Selective coding is used for model building and the generation of theoretical propositions. Coding for categories, concepts, themes, and processes took place concurrently in multiple iterations. A balanced approach to coding was considered to avoid over-coding but also to avoid under-coding. Selden (2005) warns against over-coding as a threat to creativity that can result in the codes and coding being detached from the context of the study (p. 127). This criticism by Selden may hold true if indeed the codes are designed very mechanically with very little reflection to the context. However, the constant comparison portion of grounded theory is an iterative process and the threat to creativity that Selden warns against is avoidable with careful consideration of the context in the design of the codes. There is also a threat from under-coding that will result in categories that contain more than one distinct concept and thus should be categorized separately into different categories. In

this case, categories will end up describing only at a very high-level concept, missing on the detailed distinction between similar categories and properties, and also missing the detailed distinction between similar processes. As it is described in the subsequent sections, in order to have the codes always grounded in the context, the analysis and interpretation stage is constantly informed by the memos, notes and summarization of the interviews that were written during and after the interviews, while reviewing the transcripts, as well as throughout the different stages of data analysis and interpretations.

In order to avoid over-coding and under-coding, during the open coding steps, the interviews were coded mostly at paragraph level. Some coding was also done at sentence level in instances where one paragraph contained multiple concept and categories across different sentences. Also, some codes contain multiple paragraphs such as when coding for higher level processes that were described in multiple paragraphs. To avoid under coding, whenever a paragraph contained multiple concepts that were not clearly discernible via separate sentences, the same paragraph was coded with more than one code to ensure each concept was coded.

In addition to coding for concept and categories, Strauss and Corbin (1998) emphasize the importance of coding for process as an integral aspect of coding:

“Analyzing data for process is not a separate aspect of analysis. Coding for process occurs simultaneously with coding for properties and dimensions and relationships among concepts. It is part of the axial coding and building of categories. Instead of looking for properties, one is purposefully looking at action/interaction and noting movement, sequence, and change as well as how it

evolves (changes or remains the same) in response to change in context conditions” (p. 167).

Thus, the coding for process takes place concurrently with the coding for categories and concepts.

## **Chapter 7. The emergence of four themes**

The results from the open and axial coding of the interviews are presented in this chapter. The emergent themes are described with the level of observable properties at this stage of the data analysis. These themes are further explicated in the subsequent chapters by close examination and interpretation of the quotes from the interviews with the researchers (individual and groups level) and the documentary evidence.

The open coding process (as described in section 6.4), aimed at identifying categories, concepts and processes, was conducted with guidance from Creswell (1998, p. 55-58), Strauss and Corbin (1998, p. 92-93, 101-21), and Miles and Huberman (1994, p. 55-72). The open coding resulted in 552 codes based on the 11 transcribed interviews. The open coding process also resulted in many notes and preliminary thoughts and ideas about potential emerging higher-level concept, categories and themes that can help group the main concepts emerging from the data. These notes were compared and contrasted in many iterations with the initial set of notes and memos that were taken as part of the interviewing process.

To make sense out of the 552 initial codes that resulted from the open coding of the interviews, the analysis process proceeded with the axial coding in order to identify central themes or categories that can be used to group the dynamics, experiences and processes that are enacted as researchers interact with the OA repositories. The axial coding will also make visible the dynamics, categories and concepts that are implicated in scholar's lived experience, as well as identify the corresponding properties and dimensions. The axial coding steps were guided by Creswell, Strauss and Corbin (1998, p. 123-42), and Miles and Huberman. After many iterations, categorizations, regrouping

of codes, renaming of codes, collapsing and expanding of categories and concepts, a workable picture emerged—one that revealed how the researchers experience and perceive the role and the value of the OA repositories and the access tools in their individual scholarly work as well as how they perceive the role and the value for their disciplines.

Table 6 presents the summarized result of the axial coding. A number of main categories, themes and concepts have emerged. A more complete table with the actual codes, as well as sample codes with corresponding coded texts are included in Appendix H.

Table 6: Summary of main categories, themes, and concepts and their sub-categories

Main categories	Sub-categories and/or number of codes from the open coding stage	Notes
Knowledge production process	<ul style="list-style-type: none"> <li>- Search process (99)</li> <li>- Research process (40)</li> <li>- Writing stages (44)</li> </ul>	Codes related to researchers' knowledge production process. Mostly descriptive.
Themes of discourse patterns	<ul style="list-style-type: none"> <li>- Impact on scholarly process (53)</li> <li>- Impact on scholarly output (25)</li> <li>- Integration with scholarly context (32)</li> <li>- Democratization of scholarly discourse (4)</li> </ul>	These four major themes emerge as high level central concepts with respect to which researchers discussed their lived experiences and perceptions of the role and value of OA in their knowledge production.
Value	17	A category of codes that describe the value of specific OA actors in more generic terms than those describing the four themes. Much of the content coded by the Value code is coded also under four major themes.
Role	6	A category of codes that describe the role of specific OA actors in more generic terms than those describing the four themes. Much of the content coded by the Role code is coded also under four major themes.
Access tools	44	Descriptive content of access tools.
Introduction	1	The introductory sections of each interview.
Open Access Repositories	25	Descriptive content of OA repositories.



Main categories	Sub-categories and/or number of codes from the open coding stage	Notes
Repository	8	Descriptive content of OA as well as non-OA repositories.
Data Repositories	10	Descriptive content of data repositories used by arXiv scientists.
Other, Uncategorized	144	These are the codes that covered text that contained multiple concepts ('under-coded') or text that was narrow ('over-coded'). Many of these codes overlap with the codes that have already been categorized.

Of particular interests are the four themes of discourse patterns as they represent high level groupings along the lines which the researchers describe their lived experiences as well as their perceptions about the implications of OA in their personal knowledge work as well as the implication of OA on the disciplinary information practices.

Four distinct but interrelated themes of discourse patterns have emerged:

- a) Impact on scholarly process
- b) Impact on scholarly output
- c) Integration with scholarly context, and
- d) Democratization of scholarly discourse.

### *7.1 Impact on scholarly process*

This theme is a collection of codes that reflect researchers' perceptions that the availability of the open access repositories and the access tools has had implications on their individual scholarly information practices, more specifically the knowledge production processes. Researchers perceive the OA repositories as collaboratories and disciplinary shared knowledge environments. The access to the materials therein is enhanced by the repositories' own portals, as well as by the access tools such as ADS

and Google Scholar. The researchers perceive these sets of capabilities at their disposal as enablers to find faster and easier a more comprehensive and a cutting edge research materials as open access has removed the barriers to entry both for knowledge artifacts as well as for researchers. This enables the researchers to be more productive in their scholarly work by spending less time searching across many different dispersed sources of information, and spend more time in researching and collaborating with colleagues from around the world.

## *7.2 Impact on scholarly output*

This theme is a collection of codes that reflect the perceptions about the implications of the open access repositories and the access tools on the scholarly knowledge output being produced by the researchers. Researchers perceive that the availability of pre-prints and the early availability of journal articles as post-prints in OA repositories, especially the early availability of pre-prints from many different sources and post-prints from many different journals in one place, including the availability of conference papers as well as very old articles (included by scanning), enable the researchers to identify research problems and ideas that would have been harder to identify otherwise. Thus, it has implications on the scope and the types of research problems that researchers would address at a particular time, implications on researchers' serendipitous discoveries, motivations, triggers of new ideas, and resulting in increased quality of knowledge products such as articles, class materials, conference presentations, or funding proposals.

### 7.3 *Integration with scholarly context*

This theme is a collection of codes that reflects researchers' perceptions about the role the open access tools and repositories play within the discipline. This includes perceptions about the implication of the access tools and the repositories with the disciplinary culture, perceptions about the role within the broader community of researchers in the discipline, the role of the paper based pre-print culture, the implication of the open access tools and repositories on participants' careers, etc. For example, the users of astro-ph section of arXiv, mainly astronomers and astrophysicists, perceive of arXiv as the central collaboratory where they go daily to find about the latest developments in their discipline. They perceived that the arXiv has a central role in their discipline. However, the users of PhilSci, mainly philosophers of science, perceived PhilSci as an important development in their discipline but with marginal role. They view scholars' Home Pages, invisible colleges, and JSTOR (perceived as a delayed OA resource) as augmenting PhilSci's OA role in scholars' knowledge production process.

### 7.4 *Democratization of scholarly discourse*

This theme is a collection of codes that reflects researchers' perceptions about the role the open access repositories and access tools play in the democratization of the scholarly process by enabling researchers from smaller institutions to enter the scholarly process, and also enabling knowledge artifacts such as conference papers, pre-prints that do not pass through the peer-review process, and pre-prints even before they are published in the commercial journals, to enter researchers' knowledge networks.

As patterns of discourse, these four themes encompass the rest of the categories and concepts listed in Table 6. The content from the “Knowledge production process” category and its sub-categories mostly overlaps and informs the understanding of the “Impact on scholarly process” theme. The categories that contain mostly descriptive content (“Access tools”, “Open Access repositories”, “Repository”, and “Data repositories”) inform the detailed understanding of how researchers describe and perceive the open access tools and repositories. The “Roles” and “Values” categories contain mostly researchers’ perceptions of value and overlap greatly with the four themes. The “Introduction” category contains the content from the beginning of the interviews and is helpful in understanding the career stages and researchers’ initial responses to the interviews.

To further elucidate and explain how the knowledge production process, the perceived value and role, the access tools, and the various repositories are interrelated with the four themes of discourse patterns, the axial coding process continues in Chapter 8 and Chapter 9 by analyzing the interviews with the participants, with the aim to relate “... categories to subcategories along the lines of their properties and dimensions” (Strauss & Corbin, 1998, p. 124). The detailed interpretation and analysis reveals the four themes through their specific subcategories, concepts and properties and their respective dimensions.

## **Chapter 8. Analysis and interpretation of individual arXiv participants**

To understand and develop the categories, themes and concepts that have emerged as part of the open and high level axial coding in the previous chapter, this chapter further utilized axial coding with respect to the four themes with the goal to make visible and relate their underlying properties and the properties of the sub-categories. More specifically, ANT's translation and inscription concept are used to trace, describe and interpret the relationships amongst the actors of the categories (along their properties and dimensions), and describe and interpret actors' performative abilities and the nature of the realignments that emerge and are being enacted.

The structure of this chapter is as follows. NVivo is used to extract the content for each participant related to the four themes. The analyses and interpretations proceed with understanding and interpreting the lived experiences and perceptions as expressed by each of the participants by using quotes from their interviews as evidence. Once all six arXiv participants are analyzed, the perceived properties, their relationship to the four themes and the perceived value they provide are summarized in a summary table for each scholar. In Chapter 10, all six arXiv participants are then compared and contrasted among themselves to understand the common perceptions and lived experiences as well as their differences with respect to the four themes. The same approach is followed for the PhilSci scholars in Chapter 9 and Chapter 10. The analysis and interpretation continues in Chapter 10 where the lived experiences and perceptions of the two groups of researchers are compared and contrasted at group level with respect to the four themes. Grounded theory's selective coding is applied to identify the story

line that relates and integrates the categories from the axial coding along their properties and dimensions (Creswell, 1998, p. 55-58; Strauss & Corbin, 1998, p. 143-61; Miles & Huberman, 1994, p. 55-72). ANT's methodological approach and the semantic elements are used to describe, relate and interpret the dynamics across the lived experiences and perception layer.

The analysis and interpretation of each researcher provides individual level lived experiences and perceptions related to researchers' information practices that are then used to identify the common patterns and differences between the participants within the two groups. The common patterns and differences are then used to build a group level lived experience and perception separately for the two groups and further to compare the two groups and explicate how they are similar and different and the nature of those similarities and differences with respect to their information practices. This approach to analysis is informed by Hjørland and Albrechtsen's (1995) domain-analytic approach that emphasizes on situating and understanding an individual researcher (with its locally generated knowledge and locally enacted information practices) as constitutive part in the co-construction of disciplinary domains:

The domain-analytic approach recognizes that discourse domains comprise actors, who have worldviews, individual knowledge structures, biases, subjective relevance criteria, particular cognitive styles, etc. In other words, there is an interplay between domain structures and individual knowledge, an interaction between the individual and the social level (p. 409)

Savolainen (2007) further suggests that the domain-analytic approach is congruent with the "information practice" approach in Information Science (IS) and that it can be

used to understand the mutual shaping of tools that are being developed and used individually for collaboration and sharing: “the social practice approach sees a mutually shaping relationship between information and collaboration practices and the tools developed for purposes of communication and knowledge sharing” (p. 123).

The arXiv participants are presented in the following order: A1, A3, A17, A18, A19, and A20. The non-sequential order of participant codes has been preserved for practical purposes considering the grounded theory coding was started before the recruitment process was completed. Also, as the analysis of each individual case progressed, these codes made it easier to associate the quotes and experiences with individual participants.

### *8.1 Participant A1*

Scientist A1 is an astronomer, a Professor in the Department of Physics and Astronomy at a large research university in the United States. A1 received his Ph.D. in 1984, and his name is associated with about 130 deposits in arXiv, either as an author or as a co-author. He uses arXiv and the Astrophysics Data Service (ADS) from NASA in his daily scholarly work, as well as raw data repositories such as Chandra.

#### *Impact on scholarly process (A1)*

From the perspective on impact on scholarly process, scientist A1 experiences the use of arXiv (open access repository), ADS (Astrophysics Data Systems, an access tool), and Chandra (X-Ray Observatory, as a raw data repository) as very integral to his research process to the degree that the lack of these three open access services is considered as impediment to doing research: “Well, in my own – well, in my own

research, I can't do anything without access to all three [arXiv, ADS and Chandra] of them" (A1). This suggests that the open access resources and tools have impacted A1's information practices by augmenting distinct aspects of his knowledge production process, more specifically by becoming the central means through which A1 searches for scholarly materials and raw data for his research.

"Well, it [open access] has completely changed the way communication occurs in the literature. I'm old enough to have piles of preprints as you can see some of my old piles. So it used to be that when you were at a big institution like Princeton or Harvard, you had an advantage because those places got all the recent results, all the preprints and all and so it's been much more – it's now much more egalitarian. Each – we as scientists anxiously – as we are finishing a work, we're anxiously thinking, "How quickly can I post this to astro-ph?" And what will people think? And what will people say? And often the reaction is less exciting as you hoped for, but nonetheless, it's out there." (A1)

In addition to changing the process of information searching, A1 perceives that open access has fundamentally augmented the way researchers interact with the literature in the specific discipline. This scientist perceives that the open access has removed the barriers of entry and participation for any scientist, and more importantly, it provides the scientists with the needed tools to disseminate their work as wider as possible in very short period of time. It is perceived that the speed by which pre-prints are being published has an impact on scholarly communication by enabling the scholar to distribute his work wider and faster, but also enable other researchers to have access to such knowledge much earlier that would have otherwise been possible. Instead of receiving printed copies of pre-prints and post-prints via postal mail—the pre-print culture has traditionally been an important actor in the creation and dissemination of knowledge in the astronomy and astrophysics even before the advent of networked electronic communication—the enabling networks enacted via the privilege of belonging



to the network of elite universities has been fundamentally realigned. From this perspective, the emergence of open access and its appropriation in the physics scholarly community has enabled the pre-print culture to inscribe its properties onto the digital realm of scholarly communication and translate the scholarly context for the scientists so that they can become active and participating actors in the scholarly communication landscape, more so for the scientists that were excluded from the pre-print exchange because their host institution was not part of the network of elite institutions. Thus the pre-print culture, coupled with the technical capabilities of the open access resources and tools, has realigned the scholarly communication topology by enabling new actors (scientists) to participate in the network and it has also modified at least two properties: a) time—ability to read others' works earlier and faster than what would have been otherwise possible, and distribute own work earlier and faster, and b) space—read works from a broader group of scientists that would have otherwise been possible, and contribute own work to a broader network of scientists.

“Well, it's a – I find it [open access] to be a very useful – well, it's not useful in the sense – I never go to the journals anymore. I never go to the journals. And I expect everything to be available online and so what it means is that I feel that I have access to any result that has been published and so that's a very, very powerful – it's a very powerful feeling. Sometimes it's also an overwhelming feeling.” (A1)

Here A1 very explicitly states that open access resources (arXiv and ADS) have fundamentally realigned his individual knowledge production context, where the search and research process of visiting specific journals has been replaced by the use of arXiv and ADS. This realignment is perceived so strongly by A1, that arXiv and ADS have acquired a level of trust: it is ADS and arXiv that A1 goes to find the latest results for his research interests. Thus, arXiv and ADS have been appropriated as necessary tools

modifying A1's research process as A1 searches, discovers, and accesses scholarly material.

“So, but the other thing – the thing about arXiv which – I mean it's a great idea and it too has been around for 10 or 15 years now, but I think its impact is dropping in some areas because the gap in time between publication – between submission and publication – the journals have done a great job of shrinking that gap. So, you're arXiv posting is open and available for only a short window of time before it then becomes published. And once it's published, ADS picks it up as an arXiv posting.” (A1)

A1 perceives that the impact of arXiv's translation on the scholarly publishing network of specific disciplines represented in arXiv differs depending on the time gap between the submissions of pre-prints in arXiv and the publication of the same in the respective commercial journals. This would suggest that the commercial journals are not necessarily in the periphery of the scholarly pre-print culture, rather, the commercial journals have been pulled into the open access augmented scholarly network by changing their practices to allowing submission of copyrighted materials into OA repositories. When A1 was probed during the interview about the value of open access with respect to the publication gap being reduced by the commercial journals themselves, A1 perceived that the value of open access has been diminishing with respect to the specific property of open access as enabler of articles to be read early and before they appear in commercial journals—because commercial journals have been reducing the publication gap. This seems to be a calculated response by the commercial journals—some of their processes have been translated in order to remain active actors in the knowledge production network.

Additionally, the Astrophysics Data System (ADS) as an access tool to arXiv and a digital repository of abstract and scanned older articles is perceived as an integral actor

in the scholarly communication network—very visible and functionally useable by the scientists: “And it also has links to the scanned journal if ADS has scanned it, and those are older ones” (A1). ADS is one of the actors that provides the links and relationships in the scholarly network topology, enabling scientists to easier search, discover and access arXiv submitted materials, as well as removing a barrier to scholarship by enabling scientists to access old journals in digital format. From this perspective, ADS is conduit for translation through which arXiv materials and materials from older journals that are not in arXiv enter the disciplinary knowledge network and thus are made available for use by the scientists in their individual knowledge production contexts.

“So there, the arXiv was absolutely primary to doing the project. These guys presented the positions of these four new objects and we said, “Hey, we can do a better job on the optical and x-ray follow-up.” And we went off and did it. So there, it was primus – the open access was prime in triggering that experiment – that observation – that project. And so, in other cases, it’s a race between whether it was the arXiv or when it was published in the journal and when your interest in that particular observation or that particular subject came about.” (A1)

In addition to arXiv and ADS modifying the time and space properties in the scholarly communication process, A1 perceives that these open access resources have realigned scientist’s individual knowledge network that in turn can trigger new research directions for existing projects such as “optical and x-ray follow-up”. In this case, the scientist was able to conduct research on a specific research idea much earlier than it would have otherwise been possible.

#### *Impact on scholarly output (A1)*

“No, I don’t think – I don’t make a big distinction between those [open access resources and commercial journals]. For example, this paper that I was just looking at, we started working on that immediately after we saw that paper posted to astro-ph. It was posted before it was accepted. In fact, I think our

paper will appear so our paper, which is a follow-up paper, will appear before theirs.” (A1)

This continues from the previous theme by interpreting distinct aspect of open access properties that impact scientists’ scholarly output. Although some properties clearly impact one of the themes, many of the properties impact more than one theme.

A1 perceives that the source of a specific artifact, whether found in open access repository or accessed directly through the commercial journal does not make a difference in his judgment on whether to use the article or not. Specifically in this case, the availability of the article in astro-ph, even before it was published in a commercial journal, had an effect on the article that was produced by A1 in collaboration with other researchers. It is important to note that due to the barriers of entry and dissemination that have been removed for pre-prints, as the result of the open access phenomenon, the pre-print in question had an effect on the type of research problem A1 decided to work on with his collaborators. Using the language of ANT, the open access has inscribed its properties of openness into A1’s information practices, by shrinking time to make the pre-print available for use much earlier (therefore reconfiguring A1’s knowledge network), enabling A1 to produce an artifact that otherwise would have been feasible only after pre-print’s publication in a commercial journal.

Further, in the specific cases where new research problems are triggered by reading pre-prints from astro-ph, A1 perceives that the open access availability of the pre-print was a decisive trigger in the production of new knowledge. Two properties of open access seem to have been relevant in this instance: a) the early availability of the article that triggered A1’s knowledge production, and b) the availability through a resource that enables a scientist to observe the latest and up-to-date research in their field of interest.

This has enabled the source article or pre-print to broaden the scope of scholar's research problem by enabling new and unexpected knowledge and observations to enter scholar's knowledge network.

Looking at arXiv from the perspective as a network topology of articles from many different journals that are at least linked among themselves through disciplinary belonging and their presence in a common information environment, A1 perceives that any restructuring of the topology into a number of isolated smaller networks, labeled by their narrower subject areas, will have an effect on the knowledge production by reducing the number of knowledge artifact into isolated sub-sections or sub-groups.

“No. I haven't found a real good solution to that [the overwhelming amount of materials in arXiv]. You know arXiv has come up with a subdivision of the astronomy journals – the astronomy topics. And personally, I'm not that keen on that, I'm not that happy about that. Because as bad as it is to have everything mixed together, at least there are some people who will scan the whole list and will glance over your paper as I glance over other subject areas. And if you happen to have an extra half an hour that day, maybe you will just download it and take a look at half of it. It's a challenge between trying to stay somewhat broad and trying – and being totally focused on your own subject.” (A1)

In this case, a translation of the networks into smaller networks (by putting the sectioned articles into groups) will have an impact on the type and the scope of the research problem being addressed by the scientist (due to different sets and knowledge artifacts entering the local knowledge network). There is an element of serendipity expressed by A1 that can emerge from articles from different sources that can be browsed and reviewed together in the same repository. Thus, sectioning and sub-grouping knowledge artifacts into isolated groups such as disciplinary and sub-disciplinary journal literature may realign scholar's individual knowledge production

context by acting as a barrier for the network of scholarly articles that may have otherwise entered a scientist's individual knowledge network.

“I never go to the journals. And I expect everything to be available online and so what it means is that I feel that I have access to any result that has been published and so that's a very, very powerful – it's a very powerful feeling. Sometimes it's also an overwhelming feeling.” (A1)

The availability of arXiv has fundamentally translated A1's knowledge production context where arXiv has acquired a trust as a resource of articles of interest and that it always contains the latest research findings, thus becoming a conduit for driving new knowledge production. This suggests that A1's individual knowledge production context has been realigned where A1 primarily search for the latest research findings in arXiv and ADS instead of the commercial journals.

To summarize, scientist A1 perceives that the availability of articles and pre-prints from many different sources in one common information environment such as arXiv and ADS, the early availability of the latest research, and the trust that arXiv and ADS are sufficient resources to find the latest research finding, has positioned arXiv and ADS with the ability to perform on the type and scope of research problem, and the time a specific research problem is addressed by the scientists.

#### *Integration with scholarly context (A1)*

“Even in physics, there are different cultures in terms of posting. Even in astronomy, there's different cultures about posting. There are groups of people who will post to the arXiv as soon as they submit to a journal. My tendency though is to wait until the journal has accepted the paper before submitting it to arXiv. On the other hand, though, recently I've been working more in a cosmologic – in cosmology. And there the culture, even if it's an observational paper, it's posted to the preprint server immediately.” (A1)

A1 expresses his perception about the differences in how arXiv and its sub-groups are integrated within the everyday life of the scientists. While A1 personally publishes his works after they are accepted and perceives that there are differences in when and how quickly other scientists post their research findings in arXiv, there seems to be a perception that arXiv and astro-ph are actors through which individual scientists can share their research findings immediately with the broader scientific community. The perception that arXiv and astro-ph are important actors in scientists' individual knowledge production processes positions arXiv and astro-ph with performative capability to mediate between scientists in how they search for the latest research findings: "... it's out there [in astro-ph]. It's the sort of thing that you have no excuse – you're a scientist – not to know a current piece of work. You just don't have any excuse." (A1)

Thus, A1 perceives that arXiv is incorporated in his knowledge production process and it is viewed as such by his colleagues, peers, and the larger scholarly community in astronomy and astrophysics. The perception is so strong about the role of arXiv in A1's individual knowledge production context that he perceives that there is no excuse for not being aware of the latest research. Thus, the availability of open access resources has realigned the scholarly communication topology by placing arXiv/astro-ph at the center of scientists' searching, discovering and accessing steps in his personal knowledge work.

The following quote further shows the depth of integration of arXiv and ADS in the scholarly communication culture of astrophysicists and astronomers.

"Well, I think every single experiment – every single project that I do is of that sort – that being able to sit at my desk and access all of these resources immediately without cost. Unfortunately, I don't appreciate how wonderful it is. But it's like this all the time. The reports of new targets are released. You can

get – you may not have heard about this, but you can get alerts on your blackberry from the NASA satellite telling you of the currents of the Gamma Ray Burst.

And that's open access and you can sign up for that feature. So, I'm not amazed at it any more. It's a little sad, I mean I should be more amazed, but – ... ” (A1)

A1 perceives the realigned scholarly network topology that includes the open access resources as being part of everyday life experience. A1 is not amazed any more: “So, these tools are just a natural part of what I do and I wouldn't be able to do as much as I do without them. So, I don't have the aha [moment]–” (A1).

“Absolutely” is how A1 responded to a discussion during the interview confirming that the sharing and acceptance of pre-prints has been an accepted cultural phenomenon within the broader physics scholarly community. The pre-print culture in the physics community has been in existence prior to the availability of electronic artifacts and it appears as a critical historical and a disciplinary cultural actor—as paper based open access communication of pre-prints. This perhaps explains why the culture of pre-print sharing with the larger scholarly community has been perceived as “normal” in the digital age with the availability of open access resources. From this perspective, using the language of ANT, the pre-print cultural context has been inscribed into the newly established digitally enabled open access topology, realigning the tools of communication and distribution by at least replacing the paper copies with digital copies and by replacing the postal mail distribution with the electronic distribution. The availability of the networked electronic tools have acted as conduits by which the cultural norms of the paper based open access communication of pre-prints have been inscribed into the electronic based scholarly communication in the physics discipline. In addition to replacing the medium and the tools of distribution, arXiv is a new construct



that has emerged as the shift from paper to electronic realm occurred. The emergence of arXiv as a new actor in the scholarly communication also enabled the emergence of new linking and organizational structures that would not have been feasible with the paper based exchange process.

“And so the journals still remain important. They still remain very, very important and if we did not have access, electronic access to them through records, it would be a serious impediment, a serious impediment.” (A1)

Despite the perceived central role of arXiv and ADS in the discipline, A1 perceives the journals as very important actors in the open access scholarly communication; even though A1 has stated that he no longer visits the commercial journals. The realignment in the scientist’s knowledge production context has thus been in his research process and the specific steps of searching, discovering and accessing materials needed for the production of knowledge. However, the trust for assessing articles’ quality and scholarly merit has still remained with the commercial journals due to the peer review process that is conducted by the commercial journals on behalf of the scholarly communities.

“So there are these indices that – how many papers you publish with how many citations and various different indices that were generated and it’s fairly easy, it’s actually trivial to do that with ADS. And so, we’re now in the process of the – we’re now in the – not in the process of, but we’re in the habit of assessing those things when people come up for tenure and seeing what are their most highly cited papers and who is citing them. So these provide a much less laborious way of doing citation searches – citation indices than we had in the past.” (A1)

The ubiquity and ease with which ADS can be used for citation searches, based on its perceived completeness of access to the knowledge network of astronomy and astrophysics, has enabled it to enter the dynamics of the tenure decision process where committees use ADS to assess scientists’ citation impact. ADS has thus been able to perform on the scholarly context by impacting the individual information work of

scientists at local level, that is also used by the scientists for the information work of the disciplinary scholarly functions such as the information practices of tenure committees (deciding on tenure promotions) that are enacted by its individual members but are distinct from the information work of its individual participants (writing articles).

*Democratization of the scholarly discourse (A1)*

As previously noted, A1 perceives that open access has democratized the scholarly publishing process by enabling smaller institutions to participate in the pre-print culture and more broadly in the scholarly communication that was otherwise prohibitive due to size and economic difficulties. This enablement of inclusion has realigned the disciplinary knowledge production context by enabling new actor (researchers) to participate in the networked collaboration. Further, the availability of pre-prints has realigned the disciplinary knowledge ecosystem (composed of journals and articles) by the inclusion of knowledge (from pre-prints) that otherwise would not have been available for use for the production of new knowledge. Thus, we see that the democratization principles of openness and inclusion have realigned the scholarly publishing process by enabling researchers from smaller institutions as well as independent researchers to participate in the networked collaboration. It has also enabled the opening up of the disciplinary knowledge ecosystem through which peer-reviewed articles in the form of pre-prints as well as unpublished knowledge (that did not make it through the peer-review process) can be made visible and accessible much earlier for local use by individual researchers.

*Participant A1 summary in matrix form*

The following is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A1, in relation to the four themes. In Table 7, each column represents one of the four themes. The table was built from the narrative of this sub-section, extracting and noting the properties and the perceived or experienced relationship by the scholar. The text in bold font in each cell represents a property (perceived or experienced) followed by the type of relationship with regards to the properties of various elements that comprise the personal knowledge production and knowledge network of the scholar, and the broader disciplinary and institutional context.

Table 7: Summary of OA properties and their relation to the four themes as perceived by A1

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A1</b>	<p><b>Availability</b> of arXiv, ADS, Chandra → becoming <b>integral</b> actors</p> <p><b>Openness</b> → barriers of entry removed , new <b>interactions</b> possible</p> <p><b>Time</b> (<b>read</b> earlier and quicker, <b>distribute</b> faster)</p> <p><b>Space</b> (<b>distribute</b> wider)</p> <p><b>Integration</b> → <b>reduced</b> commercial journals publication gap</p> <p><b>Linking</b> (ADS, arXiv)→ <b>enhances</b></p>	<p><b>Early access</b> → impact the production of knowledge artifacts</p> <p><b>Openness, Access to latest research</b> → <b>inclusion</b> new knowledge earlier than otherwise would be possible</p> <p><b>Sub-categories</b> of articles → might be <b>restrictive</b></p> <p><b>Inclusiveness, integration</b> of resources → enables <b>serendipity</b></p> <p><b>Trust</b> → arXiv, ADS <b>sufficient</b> to find latest research</p>	<p><b>Availability</b> of arXiv, ADS → <b>part of everyday life</b></p> <p><b>Speed, early access, inclusiveness, openness, time of distribution</b> (arXiv, ADS) → <b>central</b> actors for scholarly production process</p> <p><b>Central role</b> (arXiv, ADS) → <b>no excuse</b> not to know about the latest research</p> <p><b>Pre-print culture</b> (physics discipline) → enabled arXiv</p>	<p><b>Democratization</b> → smaller institutions and independent scientists can <b>participate</b></p> <p><b>Inclusion of pre-prints</b> that might never be published → realigned the disciplinary knowledge ecosystem</p>

	article visibility, findability, accessibility  <b>Scanned old journals</b> (in ADS) → <b>brings</b> older journals into the disciplinary knowledge ecosystem  <b>Inclusiveness</b> → <b>trusted</b> to find the latest research, triggers of new ideas and thoughts	<b>Linking</b> (ADS, arXiv) → <b>enhances</b> article visibility, findability, accessibility  <b>Scanned old</b> <b>journals</b> (in ADS) → <b>brings</b> older journals into the disciplinary knowledge ecosystem	and ADS to become <b>important</b> <b>and trusted</b>  <b>Pre-print culture</b> (physics discipline) → <b>new actors</b> <b>established</b> (inclusive, integrated, open; old/print inscribed into the new/digital)  <b>Openness,</b> <b>inclusiveness,</b> <b>integrated</b> → <b>no</b> <b>longer visit</b> the commercial journals	
--	--	--	--	--

## 8.2 Participant A3

Scientist A3 is an astronomer, a Professor in the Department of Physics and Astronomy at a large research university in the United States. A3 received his Ph.D. in 1977, and his name is associated with about 60 deposits in arXiv, either as an author or co-author. He uses arXiv and ADS in his daily scholarly work.

### *Impact on scholarly process (A3)*

“I’m an astro physicist in the Department of Physics and Astronomy here. I use the archive pre-print server at Cornell [arXiv] for posting my own work and for keeping track of the work for my peers all around the world. This is absolutely crucial for my research.” (A3)

“Well, now I think it’s absolutely essential in astronomy to post your work on this pre print server. I mean I don’t go to the library anymore. I got there for a few monographs and textbooks, but for journal articles, I don’t need to go to the library, it’s all there online.” (A3)

A3 perceives very strongly the importance of arXiv for disseminating his own work and for keeping up to date with the works of “peers all around the world”. The realignment of A3’s knowledge production process is perceived to be so strong, that

arXiv has replaced the other modes of access (such as libraries) to articles from commercial journals. A3's individual knowledge production context has been realigned to integrate arXiv as a collaboratory where the latest research results can be exchanged through pre-prints and post-prints (deposited in arXiv) with a broader group of peers.

“I go there [in arXiv] less often now. In the days before the two [arXiv, ADS] were cross linked, I used to have to go to both, but now this is cross linked into ADS, I don't need to go to the arXiv portal nearly as often. But I do still post my own contributions to the archive.” (A3)

The perception of the importance of the arXiv, as a central repository for pre-prints and post-prints, was supplemented and clarified to mean that ADS has become the central access point through which A3 searches, discovers and accesses article in arXiv. A3's has modified his research process to include ADS in a more central role, overshadowing arXiv, because ADS provides additional functionality and it also links to arXiv. As it will be explained in a later section when describing the role of the technological layer actors, the realignment has been enabled by the interconnection of arXiv and ADS at the technological level—they “speak” the same interface and the same metadata. The availability of the technologically congruent capabilities have been appropriated and realigned by the organizational layers of arXiv and ADS to enable the integration between the two. Thus, A3's individual and localized knowledge production context has been realigned both by the organizational layer (policy based enactment of cross-linking capability between arXiv and ADS) as well as by the technological layer (that makes the cross-linking possible). This distinction between the organizational and the technological layer of arXiv and ADS emerges as relevant in comparing and understanding the implications on researchers' individual knowledge production context.

“A tremendous amount of what I need, almost everything, is now accessible through these archives. They have gone back and scanned the early issues of the journals, back to the 19th Century, into the ADS system, so there’s 120 years of journal articles available online that way. It is great fun to be able to find amongst those archived, old articles, something that is really insightful from many decades ago. I’m impressed when somebody does that, and I occasionally have managed to do it myself. So yes, I guess that sort of stands out if you have a citation to an article that is from the 20’s.” (A3)

A3 perceived that ADS has enabled the researchers in the community to access articles from journals issues some of which might be 120 years old, predating the digital publications. This has been enabled by the organizational structure of ADS utilizing technological capabilities such as scanning for the conversion of paper versions of articles into digital format. In essence, ADS has performed as a translating actor and a doorway through which researchers can access very old paper based scholarly journals in digital form via a standardized interface. Thus, for this scientist ADS has become a conduit and an enabler for discovering older but insightful and valuable knowledge that can be shared with other researchers through its use in current research problems. More importantly, by bringing the older articles into the fold of the networked, integrated and inclusive information environment, contemporary researchers might make insightful discoveries and research findings, something that researchers that lacked the networked and integrated information environment that is enabled by ADS would have had hard time seeing and discovering.

“Well, the biggest challenge is that it’s so huge so that it’s hard to keep on top. As I said, there’s a lot of work every week, there’s hundreds of papers, thousands of papers a month, and so I find it hard to stay on top of that. I find the archive system is very well policed, in that, in order to post an article on there, you have to have some credibility. And generally speaking, I’m not sure if you’re aware of this, if you’re a new author who wishes to post an article on the archive, you need one or two mentors, with established reputations, who will vouch for the fact, yes, this person, and they will not allow somebody from aol.com to just put their crazy idea into the archive. That is excluded.” (A3)

A3 perceives that the organizational structure of arXiv is well managed in order to balance between a large amount of data being available and its usefulness for the scholarly community. The process of “reputation establishment” for newcomers via the vouching of two existing members of the community acts as the keeper of the balance regarding the relevance of submissions without making any judgment on the quality of a specific work. The emergence and the establishment of “reputation” with respect to the actor and process is intrinsic to the new network topology—augmenting the process that was carried by the handful of institutions that participated in the pre-print circulation culture. This is an example of realignment of the production process by way of a new actor (i.e., reputation) emerging within the new context—albeit there are some properties of the peer-review process inscribed into it, translated on the way as a moderation process to enable possible peers to enter the scholarly network, instead of “moderating-reviewing” the quality of an article for entry into the disciplinary knowledge ecosystem.

*Impact on scholarly output (A3)*

“Generally speaking, I’d prefer to cite papers which are either in the journals, or accepted for journals. There’s a strong preference to do that because these are articles which people you know, some referee has blessed. I feel capable of forming my own judgment, but it is the case. I write articles sometimes for conference proceedings in particular, in which I will put original material [that] never appears anywhere else. And so that material, I hope people will cite, and it is cited, I have noticed that my work is cited in that. But it is the case that if I publish work in a journal article, it is much more cited. So that is important for my field, and I behave the same way when I cite. I try to find the article in the journal.” (A3)

With this statement, A3 delineates the information resources that define the disciplinary knowledge network into those in the center (such as journals) and those in

the periphery (such as conferences and conference papers). This seems to suggest that the knowledge network of a discipline, as it is manifested through the materials published through the information environments in the center, can be realigned by the inclusion of knowledge from information environments in the periphery. This presents another notion of realignment of the disciplinary knowledge network where the open access resources from the center and the periphery are aggregated together and presented to the researchers through an integrated information environment. Vetting the quality of the materials obtained is still a job that researchers have to perform. Nevertheless, with open access, conference proceedings, pre-prints and post-prints are all searchable and accessible via the same normalized process (same browsing and search interface; a search query returns different knowledge artifacts at the same time) and all have the performative ability to enter the individual knowledge networks of scientists for use in articles, conference proceedings and other forms of scholarly output.

A3 perceives that the various articles he finds in arXiv or via ADS often are triggers for new ideas that lead the scholar to produce new articles: “Most of the time it’s an article [that triggers new idea]” (A3). This would suggest that A3’s individual and local information practices and knowledge network has been realigned to incorporate the open source resources as inclusive actors with their performative power to initiate the realigning of scholar’s individual knowledge production context, enabling the scholar to produce articles using a different subset of knowledge artifacts.

*Integration with scholarly context (A3)*

“In astronomy, still we have refereeing processes in journals, and you might post your paper sometimes before the refereeing stage, but usually afterwards, if it



hasn't previously been posted on the archive as well as appearing in the journals." (A3)

A3 perceives that the process of posting to arXiv has been integrated within the disciplinary production network topology alongside the peer-review process performed by the commercial journals. Posting to arXiv as an accepted practice can happen before the pre-print has been accepted for publication, after it has been accepted for publication or sometimes after it has been published.

"And it's certainly the case in the astronomy world, the papers which are ultimately accepted to appear in journals are far more referenced than those which are posted but never appear in journals ... And those generally don't have nearly the impact of something that's refereed and appears. " (A3)

"But it is the case that if I publish work in a journal article, it is much more cited. So that is important for my field, and I behave the same way when I cite. I try to find the article in the journal. I suppose one still feels that's more accessible but the different between the two is not, it is very marginal. " (A3)

Here we see that A3 perceived that the submission into arXiv enhances the chances for an article to be cited after its publication. Publication in commercial journals is still the most critical (reputation networks is still enacted and managed by journal systems), however, once the article is published and it is deposited in arXiv either as pre-print or post-print, its inclusion in arXiv and its subsequent availability through ADS increases its visibility, therefore these article receive more references. Thus, the inclusion process of articles in arXiv and ADS instigates a performative process by which the articles in arXiv appropriate an enhanced property of visibility that in turn translates into performative capability through articles' inclusion into scientists' knowledge networks. Article's performative capabilities are twofold: a) because of the increased visibility, the article will be read by more researchers and therefore an increased chance for more

references, b) the knowledge represented in the article will impact the structure of the article in which it is used.

“I have not used Google Scholar. I don’t use any – maybe I’m backward but I – This [process of using arXiv and ADS] fits my needs so perfectly that I don’t really –” (A3)

This very strong statement means that A3’s use of arXiv and ADS provide a complete access to the open access resources available in this discipline, and that generic access tools such as Google Scholar are not seen as relevant in the context A3’s information work. The performative aspects of ADS have enabled it to be appropriated in the scholarly discipline as a critical actor becoming the most relevant access tool for astronomers and astrophysicists. The appropriation by which ADS has been incorporated into the knowledge production of astronomers and astrophysicists has been enabled by the technical level capabilities that ADS is composed from (such as the cross linking between ADS and arXiv), as well as the organizational structures that have linked ADS to arXiv and other abstracting services.

“They still are, I mean Springer articles I know I can’t get to through the ADS system, and there are one or two other publishers who, they request a credit card at some point to be able to read an article, which I don’t do. But since so many of those articles are actually on the pre print server, if that’s the case, since they’re cross linked in ADS, I just go to the pre print. ” (A3)

Here A3 indicates that although some commercial journals that are important and still not accessible by ADS, their articles can be found in arXiv as pre-prints. It is important to note how the availability of arXiv and ADS have realigned A3’s knowledge production process by their inclusion where for the most part commercial journals have been excluded as direct actors in the production process, where arXiv and ADS can be seen as aggregating proxies through which researchers access commercial journal

articles. However, as A3 has stated earlier, commercial journals are still very critical for the peer-review process. Thus, although the commercial journals have been excluded and no longer belong in scientists' direct search process, the commercial journals are a critical element in scientists' knowledge production process through the performative aspects of the peer-review process.

“No journal that I know of is missing. Everything that I want to find – Well let's think, there are articles in Russian which are published in English translation, and when they're published in English translation there's a significant amount of work done in writing Russian these days in my area. So maybe yes, but I don't know what value it would be if they were in Russian. ” (A3)

Scientist A3 perceives that the open access resources in astronomy and astrophysics are complete and do not miss any articles, except for articles published in other languages, more specifically in Russian. In the language of ANT, it would mean that the astronomy and astrophysics knowledge ecosystem is almost completely defined and encompassed by the open access resources. The same resources are available by other access venues as well. However, A3's perception positions the relevant knowledge ecosystems represented by arXiv and ADS as sufficient and complete for participation of scientists in the relevant disciplinary scholarly context.

Finally, scientist A3 perceives that arXiv is the central hub through which research findings are distributed and it is the primary conduit that provides the necessary community visibility to research materials: “And I think that's a general recognition amongst the community, that if you don't post your article on archive, then you may as well not publish” (A3). This perception associates strong central performative role for arXiv that is recognized by individual scientists as well their research communities.

*Democratization of the scholarly discourse (A3)*

“One thing I particularly like about this archive system is that it’s extraordinarily democratic, that is that you don’t have to be at a major research university to get access to it. When I went through my training and post-doc era, you had to be at a major research institute in order to be able to see the pre prints as people use to mail out through the airmail, and hard copies of papers to the major institutes in the world. So you typically send out 300 or 400 of these copies of your article to all these different institutes, and if you were not at one of those institutes, you never got to see what was currently patenting in your area until the articles appeared.” (A3)

A3 also perceives that the availability of articles in open access repositories is democratic in the sense that it enables scientists from all over the world to participate in scholarly discourse even if they do not have the means to subscribe to expensive commercial journals. This is a continuation of the pre-prints culture that was exclusionary in nature—only well-established universities (limited in number due to cost and economics) could receive pre-prints in actual paper print copies. The conjunction of the electronic networked information environment and the open access phenomenon have thus enabled the inclusion of actors (scientists and institutions) that otherwise could not participate in the network.

In response to a final prompt at the end of the interview whether he would like to add anything else, A3 responded:

“I don’t think so. I think I’ve told you – I mean what I really wanted to say was that thing about democratization. I think that was great. I mean I am very impressed with that, and we’re relatively new at [A3’s institution] in astronomy. It’s not an established place where astronomy has been done for 60 or 70 years. We started only 30 years ago, and we would never had been in that pre print circuit. So from that point of view, it’s vital for starting a new research group in the area.” (A3)

Again, he strongly emphasizes the democratization property of the open access resources. It appears that in addition to democratizing the context by enabling the

inclusion of smaller and less established universities, it has also enabled the establishment of a relevant research group at a major research institution that otherwise would not have been able to emerge and become part of the disciplinary knowledge production network topology of astronomers (such as the astronomy group at the Department of Physics and Astronomy at A3's institution). Here we see that the open access has not only realigned the knowledge production network of a specific discipline (by enabling researchers from smaller institution to enter the disciplinary discourse), it has also helped the group of researchers to establish themselves locally with their institution by using global resources such as arXiv and ADS.

*Participant A3 summary in matrix form*

As it has been already described, Table 8 is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A3, in relation to the four themes.

Table 8: Summary of OA properties and their relation to the four themes as perceived by A3

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A3</b>	<p><b>Openness, inclusiveness, integration → Collaboration</b></p> <p><b>Linked, interconnected</b> (ADS, to arXiv) → arXiv not visited as much</p> <p><b>Inclusiveness</b> (of sources that have not gone through peer-review)</p> <p><b>Inclusion</b> (access to</p>	<p><b>Inclusiveness, integration</b> (arXiv, ADS) → trigger <b>new ideas</b> and research</p> <p><b>Inclusiveness</b> (of sources that have not gone through peer-review)</p> <p><b>Inclusion</b> (access to journals article from 120 <b>years ago</b> or so, scanned)</p>	<p><b>Read and deposit</b> pre-prints (arXiv) → include arXiv and ADS as <b>integral actors alongside the peer-review</b> process</p> <p><b>Deposit</b> in arXiv → citation impact <b>increases</b></p> <p><b>Trust, Total access to all relevant resources</b> via</p>	<p><b>Inclusiveness</b> → Democratization by inclusion of smaller and less economically well institutions</p> <p><b>Inclusiveness</b> → smaller departments from established institutions can also participate or be established</p>

	journals article from 120 <b>years ago</b> or so, scanned)  <b>Reputation  establishment</b> (arXiv) → <b>balanced and  relevant inclusion</b> of materials in arXiv		arXiv, ADS → <b>commercial  journals  marginalized</b>  <b>Total access to all  relevant resources</b> via arXiv, ADS → <b>good enough  proxies</b> that define the disciplinary knowledge ecosystem	
--	--	--	--	--

### 8.3 Participant A17

Scientist A17 is an Assistant Professor of astronomy and astrophysics in the Department of Physics and Astronomy at a large research university in the United States. A17 received his Ph.D. in 1998, and his name is associated with about 100 deposits in arXiv, either as an author or as a co-author. He uses arXiv and ADS in his daily scholarly work, with occasional use of Google Scholar.

#### *Impact on scholarly process (A17)*

“I think everybody agrees that archive plays an important role and uses that I think pretty much across the field, at least that’s my understanding from talking to other people, that everybody uses that as the first source. If you wanna keep up with what’s going on in astronomy, you look at the archive.” (A17)

A17 perceives that all astronomers use arXiv as the main source to keep up-to-date with the latest research findings in the field, echoing the theme of impact on scholarly process and qualifying it as a general agreement amongst the astronomers. With the language of ANT, arXiv is the link through which the realignment of the knowledge ecosystem in astronomy can be viewed and consumed. Even more specifically, arXiv is the source where the new actors (knowledge artifacts) that realign the knowledge ecosystem of astronomy first appear.

“I imagine that the role that archive plays in astrophysics I would assume is similar to the role that it plays in other disciplines. I don’t know too much about other disciplines, but I guess for us it has become the principle source of disseminating results.” (A17)

In addition to arXiv being used as a source for accessing the latest research findings, here A17 emphasizes that arXiv is a conduit for dissemination of the latest results. This places arXiv as an important actor in the scholarly communication of astronomers where arXiv has performative power to mediate in how research findings are shared, communicated and consumed. More than a shared information resource, arXiv is an actor through which the locally produced knowledge enters the disciplinary knowledge network.

“I think it definitely has an impact because if nothing else, I imagine it has accelerated the rate of the discussion because we no longer have to wait the six months or a year or whatever it would take to go from being finished with a piece of work to actually having it appear in the journal such that it would be distributed to other people.” (A17)

A17 further elaborates on the type of realignments that have occurred in the disciplinary knowledge production network and the disciplinary knowledge ecosystem. New knowledge (research findings, discoveries, etc.) enters the knowledge ecosystem much earlier, as a result of which an accelerated knowledge exchange will occur that can actually have an impact of what other research findings can enter the knowledge ecosystem. Thus, the availability of arXiv has realigned both the personal knowledge production process as well as the disciplinary knowledge ecosystem by accelerating the circulation of knowledge artifacts as carriers of the “accelerated the rate of the discussion”—with the properties of time and “newness” emerging as properties of the newly enacted dynamics in A17’s individual knowledge production context.

“I guess one thing that’s nice about ADS is because it’s out there and so I can access it from my work desktop, from my laptop, from my home computer or whatever, as opposed to trying to remember I know I downloaded that paper, but which computer is it on.” (A17)

A17 perceives ADS as an access tool that is location independent. As a means to finding and discovering knowledge artifacts in the knowledge ecosystem of astronomy, it provides somewhat a uniform process to searching articles independently of their original source. This indeed streamlines the process by enabling scientists to always search via the same approach across the numerous article as well as abstracting sources. Thus, it has inscribed its uniform approach to searching into scholar’s information search practice by translating it into an experience that does not change from source to source, even though the different journals and abstracting service represent different but complementary knowledge networks and most probably have different interfaces. By being location neutral, arXiv has also become the “personal” and “localized” archive environment for individual researchers, even though it is a disciplinary level information environment: it is global and yet it is very local, “I do not keep it [pdf file of an article] because I’ve found that it’s usually easier for me to find the paper again online” (A17), augmenting scientists’ work by removing location dependence in the process of finding and managing scholarly work from an office, a laptop, or home office.

“So depending on how you want to interpret – we’ve been talking principally about literature archive. There are some other tools that I have just begun to use, so for example Subversion, SVN, as a repository for documents. This came up when I was working on a project with two people at the Institute for Advanced Study.

We’ve decided that the best way to keep a repository of the documents we were working on for this project was to produce an SVN repository to which we could upload and download documents as a way to exchange documents rather than e-mailing them back and forth.” (A17)



Perhaps the SVN can be viewed as a local and temporary extraction of a portion of the arXiv repository that contains the documents that enable easier and more efficient exchange of ideas and thoughts while working towards the same goals with collaborators. In this case, arXiv has enabled the emergence of a temporary structure for the duration of a task (writing an article). The emergence of the temporary structure, the workplace, is not necessarily dependant on arXiv, however, arXiv and ADS have performed on A17's local knowledge network by streamlining the selection process of the initial materials that are deposited in the temporary localized workplace: "... so it's basically an environment for collaborative work. That I mention only because it was something that is more localized." (A17)

*Impact on scholarly output (A17)*

"I would say one episode that comes to mind is more an issue of serendipity. This was several years ago, within the span of – I don't remember if it was within the same day or just a few days apart, there were a couple of papers related to my research field that struck me as interesting just before I was getting ready to go visit a collaborator. So I just took those papers, and I printed them in that case, and I took them with me, and said, "These two look interesting. They were very different from each other, but they related to both of our interests," and they led to a series of three papers that we wrote together." (A17)

This is interesting case of serendipitous discovery due to the ability to see papers in one spot from different sources related to different subjects. In this case, A17's individual and local knowledge network was reconfigured through an inter-disciplinary collaboration with another researcher that produced a set of articles that otherwise would have been a challenge to produce.

"That was the benefit of looking at the daily update. There were a couple of papers each of which caught my attention separately, and then I sort of said,

“Wait a minute. I can kind of tie them together via my interests and my collaborator’s interests in a way that I think other people hadn’t thought about before us.” (A17)

The daily update emerges as a trigger for the realignment of A17’s individual knowledge network. But before the daily update could trigger such reconfiguration, its capabilities and functions had to be technically built and appropriated by the scholar. Therefore, the realignment and impact on scholarly output (different in content, not necessarily better) results from an amalgamate of performative capabilities at three different layers: technical (the software that enables the daily updates features such as “new”, “recent”, “current month’s”), organizational (the organizing structures that appropriated the technology and enabled the various updates as feature), and the scholar who appropriated the capability in his scholarly information practices, more specifically the scholarly search process. The visibility and interconnectedness of the articles that were enhanced by the organizational and technological capability of arXiv have in this case acted as enablers of social level collaboration.

“The daily update I would say more akin to simply discovering new ideas. I’m a little bit probably more likely to discover things just in the raw discovery sense that way because I’m looking trying to look at the titles and abstracts of all the papers that have appeared. Astronomy is growing but still manageable to look at all of them. You cannot read all of them, but at least look at all the titles and abstracts.” (A17)

A17 perceives that the daily updates mechanism available via arXiv instigates new ideas for research. The flexibility and the availability of many different sources in one place has translated arXiv into a comprehensive source bringing new knowledge (through their representation in articles) in close proximity to each other. Therefore, patterns of knowledge structures represented in multiple articles can be easier to observe

(“simply discovering new ideas”), and can lead to new research (“discover things just in the raw discovery sense”) and new knowledge artifacts being produced.

“I would say that that’s the benefit of having everything in one place so that these two articles that would probably have gone into different journals – they had somewhat different target audiences, but because there’s a single astrophysics archive, they both appeared.” (A17)

“And the series of three papers that resulted from that moment were a little outside of my mainstream research, so it was a case where, “Aha, this looks interesting.” It sort of took me in another direction for a little while. Not that far away, but still unless I had seen these two papers together, I’m not sure I would have gone off in that direction for a little while.

That is a case where I do feel as though had I had any sort of filtering system in place, then I’m not sure I would have seen them and done that little scholarly diversion for a while.” (A17)

In this instance A17 perceives that arXiv, with its property to trigger serendipitous discovery, has realigned his research interests by introducing him to tangential but relevant area of research interest that directly resulted in A17 collaborating with other scientists to produce three papers. A17 directly emphasizes and recognizes that the integration property of arXiv and ADS were directly responsible for the production of his knowledge artifacts. The ability to filter out materials into different sections or bring materials together is an emergent property with immediate and direct implications performing on scientist’s collaboration efforts (social realignment) and the type of research problem being investigated and worked on (epistemic realignment). Any potential filtering systems or framework would have performed on A17 knowledge network as a barrier and inhibitor to serendipitous discovery. Thus, the performative capabilities have been enabled through arXiv’s properties of openness and integration of disparate resources that have inscribed themselves into scholar’s knowledge production process, being translated on the way as a knowledge artifact.

“At this point, what the archive structure is, it’s a database – if the database gets too large then it will fracture into several subfields, and I suppose I’ve been – I think it’s good that astronomy has not splintered yet, as I mentioned. I suppose it’s worth thinking about whether there is some other structure that would allow you to cut down on the volume without impairing your ability to find things that are a little bit outside your zone of familiarity to have the serendipity. I haven’t thought much about what you would do along those lines.” (A17)

A17 values greatly the serendipitous possibility enabled by arXiv and ADS and views any attempts to sub-divide the astronomy and astrophysics sections within arXiv as a barrier to learning and discovery.

*Integration with scholarly context (A17)*

“So historically, that was referred to as preprints, prints of a paper that were in advance of publication. People would actually print out copies and physically mail them to collaborators. So that is what is done electronically now.” (A17)

A17 perceives that arXiv is an electronic manifestation of the historical pre-print culture that has been intrinsic to the physics community for the last hundred years or so. From the perspective of the pre-print-culture in the physics community, they have appropriated the features of electronic information processing efficiency and electronic networked collaboration to achieve what has always been part of the physics scholarly culture. However, transposing the paper and postal mail practice into the electronic networked environment has opened up new possibilities that were not feasible with paper and postal mail.

Thus, the migration of the “old” to the “new” enabled the emergence of new structures and tools such as ADS, citation linking back and forth, more efficient abstracting and aggregated search, etc. This is manifested through the realignment of various actors at three different levels (technical, such as the software and systems-to-systems interfaces that enabled arXiv and ADS; organizational, such as ADS; and lived

experiences about the perceived changes in the search process by using ADS) as well as the knowledge ecosystem of astronomy and astrophysics—in essence realigning the knowledge production context of astronomy and astrophysics.

“In fact, for many people it [arXiv] plays the role traditionally played by the journal and has actually sort of eclipsed the journal in terms of the resource that people use most, I think.” (A17)

This is a very strong statement by A17 to perceive that arXiv has replaced the traditional role that commercial journals played as points of individual convergence on per journal basis where you contribute new knowledge. Certainly, the commercial journals still play a major role in the peer-review process, but arXiv is the place most researchers visit most often to search, discover and access new articles.

“Astronomy, we have a modest number of different journals, enough that you can check what’s in the Astrophysical Journal, what’s in Monthly Notices of the Royal Astronomical Society. You can sort of keep up with the different journals. My impression is that’s not true of other fields. In mathematics for example, there might be many different journals which have different areas of specialization. So there may be a little bit more compartmentalization of knowledge and of results there, whereas again for us archive [arXiv] really provides a single repository that almost everybody in the field is familiar with” (A17)

A17 perceives that arXiv is very central to the scholarly knowledge production process in astronomy and astrophysics as everybody knows and uses arXiv as sufficient resources to keep up to date with the latest research. Whereas before arXiv the pre-print culture was a paper based distributed information exchange process, arXiv emerges as a central and common structure whose properties of open access, integration and inclusiveness enhances collaboration and exchange of knowledge for further use. As few other arXiv scientists have expressed themselves, A17 relates his personal experiences as being part of the broader experiences of other astronomers and astrophysicists—

positioning arXiv and ADS as the topologies of reconfiguration that link the individual and local experiences with experiences of the disciplinary level perceived to be experiences by the other scientists.

*Democratization of the scholarly discourse (A17)*

“My understanding is that archive has played a more vital role – the open access aspect of archive has been more important for people at smaller institutions that may not have subscriptions everywhere or in particular for people around the world that may not have direct access to subscriptions. So one thing that’s nice about archive is that it is entirely open access.” (A17)

A17 perceived that the main property of arXiv, its open access policy, has inscribed itself into the disciplinary scholarly knowledge process by making it more porous and thus enabling the participation of new actors (scientists and institutions) in the scholarly discourse network. A less obvious entrant into the knowledge production context is the new knowledge that has entered as a result of the openness and integration. This is the knowledge that has been produced by serendipity and the availability of various resources in one place in integrated way, as well as by the new knowledge that has been made available as pre-prints that would otherwise not be available if it did not make it through the peer-review process.

*Participant A17 summary in matrix form*

As it has been already described, Table 9 is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A17, in relation to the four themes.

Table 9: Summary of OA properties and their relation to the four themes as perceived by A17

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A17</b>	<p><b>New link</b>, arXiv as a Channel for <b>dissemination</b> of research results</p> <p><b>Time and space</b> (quicker and wider dissemination of research results) → Knowledge enters disciplinary knowledge ecosystem much earlier</p> <p><b>Time</b> (Early access to latest research results) → realignment of knowledge search process</p> <p><b>Space/location</b> independent → enables searching from any location</p> <p>Bring <b>uniformity</b> to the search process (ADS, arXiv) → many different sources can be accessed with the <b>same interface</b></p> <p><b>Openness, integration, close proximity</b> → triggers <b>new ideas</b> and research projects</p> <p><b>Space, time, integration</b> → enable <b>collaboration</b></p> <p><b>Openness, integration/space</b> (arXiv) → enables serendipity</p>	<p><b>Openness, integration/space</b> (arXiv) → enables <b>serendipitous ideas</b></p> <p><b>Space/time</b> (daily updates) → <b>trigger</b> for new ideas</p> <p><b>Sub-categorization</b> (arXiv) → <b>barrier to learning</b> and discovery</p> <p><b>Time</b> (Early access to latest research results) → realignment of <b>knowledge network</b></p>	<p>Old pre-print print based culture <b>transposed</b> in the digital realm (knowledge production context <b>realigned</b>; <b>new entrants</b>)</p> <p><b>Complete, trusted</b> (ADS, arXiv) → has <b>replaced the commercial journal</b></p> <p><b>Very visible, open, central, common structure, integrated</b> (all research in astronomy/astrophysics know about arXiv/ADS) → has <b>realigned</b> the disciplinary knowledge production context</p>	<p><b>Opens, porous network</b> → <b>new actors</b> such as scientists, institutions, access tools, etc. <b>can participate</b> in the epistemic culture (realignment of the knowledge production context)</p>

#### 8.4 Participant A18

Scientist A18 is an Assistant Director in the department of Office Operations at a large observatory in the United States. At the time of the interview she was in the later stage of her Ph.D. studies and has subsequently finished her Ph.D. in Technology Management. She has primarily used arXiv and CiteSeer (as access tool) in her professional work and for software development. She is an interdisciplinary researcher that has used academic artifacts in her line of work related to information management and information quality.

##### *Impact on scholarly process (A18)*

“I think they’re [open access] much more important because I think that through open access you have the ability to get research that – there’s two things that you get to do. Either you get to bypass the one-year plus time span between the author submitting a publication and you actually getting to read it, which is really nice. The second thing is that certain things that ordinarily would not be considered mainstream, they might not get into the mainstream journals, you have access to them, and that’s very important.” (A18)

With respect to open access’s value in scholarly process A18 related the importance of open access in its enablement of scientists to access articles in pre-print stage and therefore gain access to new knowledge that otherwise would take a year to get to (reducing the publication gap), and the ability to read research and findings that either are not mainstream or would not make it through the peer-review process and therefore would not be published. We see that the open access principles as manifested through arXiv, impact the rate to which new knowledge (as represented in knowledge artifacts) is consumed and used for further research, modifying how scientists search, discover and access the article. Thus, arXiv acts as the conduit through which the property of openness is translated into the unrestricted access to articles for the scientists. Second,



the availability of knowledge in pre-prints (those that do not make it to the commercial journals for publication) has the ability to realign scholar's localized and individual knowledge network by bringing in unpublished scholarly materials that in turn can trigger ideas for research that would otherwise not be possible.

“Well, I mean, when you submit something to a journal or any sort of publication you have your editors and your reviewers. The editors are looking to find articles that they feel will be interesting to their readership. It is one person or group of others making a decision on relevance for an entire readership. Then you also have your reviewers who it's in their interest to criticize, so their job is to find areas in which your article needs to be revised or reviewed or perhaps to call out what they believe are errors in your methodology or your approach. As a result, there's a bias there based on the knowledge of the reviewers.

I think it's unfortunate that in a lot of cases it's either a shining result from maybe four reviewers, but then you have a very bad result from one reviewer and it still stops you from getting your article published. Even when in fact it's three years later and you find out that you really did have a good article and it was just that one person blocking it from going in, so we really need to move beyond those sorts of barriers. I think open access is the first step in helping us do that.” (A18)

A18 perceives that the peer-review process is a barrier to entry for new knowledge, due to the article review dynamics between editors and reviewers with their gate-keeping agency, and that the open access principles as inscribed in arXiv have appropriated mechanism to enable quality knowledge to enter the disciplinary knowledge ecosystem. Thus, arXiv is perceived as translating actor that performs alongside the peer-review regime, supplementing the sharing and discovery of knowledge artifacts amongst the scientists.

“It [publishing in commercial journals vs. open access venues] would change my use of the resources, but my publishing didn't change my use of the resources. I think what's changing is my perception of, like, when I have some research being done or a book chapter is being done I want other people to be able to find it. So I think from now on I'm gonna be a little more discriminating about where I choose to put my work, so I would seek out things that are more open access now than before.” (A18)

A18 perceives that the open access enables better visibility of published materials when compared to materials that are just published in commercial journals. This perception has reconfigured A18's publication approach by consciously and selectively planning to publish in open access venues in the future.

"I've published a lot in commercial publications and I think it's a great way to get started meeting collaborators, but I'm really disappointed that it's so hard to find those documents. I would love to do more open searches of things that have been commercially sold, but they're just not out there. So one of the worries that I've had lately is oh my goodness, I'm producing all this stuff that really is useful and it's helped me find communities of collaborators, but nobody else can find it. I can't even find it myself when I look and I know exactly what I'm looking for. I think that what's that done for me, and this is just briefly over the past month or so, is that I think I'm gonna start working less on commercial publications." (A18)

A18 also perceives that the ecosystems of knowledge artifacts available exclusively in commercial journals are not easily searchable, discoverable and accessible; i.e., the isolationist nature of commercial journals from each other is a barrier to knowledge discovery and sharing. Therefore, with open access enhancing and realigning scientists' search process to enable easier and comprehensive access to scholarly artifacts, it creates the conditions in the network topology for actors such as access tools and personal search practices to enhance scientists research processes.

"I value the process the most because I'd be willing to deal with issues in quality. Clearly, when you're reviewing an article, you can tell whether or not it's high quality. You can kind of search through and distinguish those yourself." (A18)

Recognizing the problem posed by the lack of peer-review of pre-prints and that some might never make it through the peer-review process, A18 perceives that established scientists that are more advanced in their career can distinguish the quality of the materials used for their research and decide whether to use an article or not.

*Impact on scholarly output (A18)*

“I use them [arXiv, CiteSeer] very, very frequently in the beginning part of doing research. I use them extensively for literature reviews because through our library at the office, you can get to a lot of different journals, but some of them you can’t get to without the preprint service.” (A18)

A18 uses arXiv and CiteSeer mostly early in the research process to explore the knowledge ecosystem for specific research problems. And even though A18 is part of a well-established research organization, her research output will lack completeness because not all of the resources she needs for her research are available through her institutional membership. Thus, expanding A18’s network of knowledge resources she can use.

“The other thing that I find really frustrating – this kind of goes back to a comment that I said earlier. That, say for example, in Archive [arXiv] you have to choose between comp-ph and the different disciplines, so that really restricts your ability to find the information you’re looking for.” (A18)

“I mean, for example, if you can deal with it now because there is – you can see all the different groupings on one screen, so I can imagine 20 years from now if there’s hundreds of different groupings that that would be really challenging.” (A18)

A18 perceived that arXiv is restrictive in the sense that for scientists that are interdisciplinary, such as herself, and need to search different sub-section of arXiv, the search process gets prolonged unnecessarily: “It just draws out the process. That’s all” (A18). This is an example of how disciplinary level epistemic requirements for knowledge production (for interdisciplinary researchers vs. researchers bound by a single discipline) have shaped scientist’s perception about the value of sub-categorization and filtering as either a barrier or an enabler to knowledge discovery. For interdisciplinary scientists such as A18 the sub-categorization is viewed as a challenge and a barrier with restrictive performative power in a sense that will inhibit serendipitous

discovery (needs to enact multiple searches in different sub-sections), while for scientists that are bound by a single discipline with a corresponding section in arXiv such partitioning is welcomed as it reduces information overload. This would seem like a reversal of the realignment that arXiv has enabled the scientists by integrating multiple sources to be accessible through one access point such as arXiv and CiteSeer. It reveals that the disciplinary knowledge networks and sub-networks are dynamic in nature based on the epistemic requirements of the discipline and its practitioners.

“I do think that the open access repositories are – you should make them more, and I do use them more when I’m building proposals because you do have access to things that are a little more cutting edge that perhaps the mainstream journals wouldn’t take.” (A18)

Here A18 emphasizes that open access, through its enablement of cutting edge research that does not get published in the commercial journals (as carriers of normative knowledge), has an impact on the scientists’ knowledge output by entering their the disciplinary knowledge ecosystem of the discipline.

#### *Integration with scholarly context (A18)*

“You know, I don’t think I’ve had a wow moment because I’ve been using them for so long. If I had one, it was probably years ago, and I can’t remember what it was like.” (A18) is how scientist A18 responds to the interview question about whether the open access resources have been extremely helpful, or if they have produced a “wow” moment in her scholarly work. A18 response suggests that arXiv and CiteSeer have been so much integrated with her scholarly context that there are no moments of surprise or excitement. For A18, arXiv and CiteSeer are part of how she interacts with the scientists

and the knowledge artifacts in the discipline: “It is definitely part of my standard operating procedure and has been for a while.” (A18)

*Democratization of the scholarly discourse (A18)*

“The way that it’s perceived is that a day-to-day like Archive [arXiv] that presents you with preprints and post prints is absolutely essential for your research... I’m sure for other people who are in smaller universities that might be even a bigger issue than what I have.” (A18)

A18 perceives that the broader community’s access to latest research via pre-prints and post-prints from arXiv is very critical for scientists’ research. This has enabled scientists from smaller institutions and the institutions themselves to participate in the scholarly discourse and be able to reconfigure the disciplinary knowledge production context by becoming integral actors. In addition to the democratization principle of enabling new entrants (such as scientists, departments, and institutions) into the scholarly discourse, A18 perceives that open access as enacted through arXiv has enabled unpublished knowledge to enter the disciplinary knowledge ecosystem.

What A18 values the most is the perception that the democratization principle as a property of the open access phenomenon in scholarly publishing not only enables new entrants into the scholarly production topology, it also democratized the production of scientific knowledge by realigning the process to not necessarily depend on the peer-review regimes, at least for more experienced and advanced researchers: “Quantity is okay, but I think the parts that I value the most is how this has democratized the production of science resources” (A18).

*Participant A18 summary in matrix form*

As it has been already described, Table 10 is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A18, in relation to the four themes.

Table 10: Summary of OA properties and their relation to the four themes as perceived by A18

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A18</b>	<p><b>Time (Early access)</b></p> <p><b>Access</b> (to materials that otherwise would not be available)</p> <p><b>Time, access</b> → enhance the <b>rate of research</b> (research process realigned--find, discover, access)</p> <p><b>Openness</b> → enables <b>new knowledge to enter</b> the disciplinary knowledge ecosystem (this is knowledge that didn't get through the peer-review process)</p> <p><b>Openness</b> → pushing the scientists to <b>publish more in open access</b> venues</p> <p><b>Openness</b> → enables <b>article visibility, discoverability, access</b></p>	<p><b>Openness, completeness</b> → enables <b>more complete and well informed research</b></p> <p><b>Time, space, openness (early access)</b> → more <b>relevant and cutting edge research possible</b></p> <p><b>Time, access</b> → enhance the <b>rate of research</b> (pre-prints that don't make it through the peer-review do trigger ideas and thoughts)</p> <p><b>Inclusiveness, openness</b> (pre-prints) → researcher would <b>consider and incorporate knowledge that is not necessarily mainstream</b>, or has not been published in commercial journals</p> <p>arXiv <b>sub-categories / categorization</b> → <b>prolong the time</b> in the research process for <b>interdisciplinary</b> scientists; <b>inhibits serendipitous discovery</b></p>	<p><b>Trusted, complete</b> → integrated with scholar's everyday scholarly endeavors</p>	<p><b>Openness, inclusiveness, integration</b> → <b>democratized</b> the production of scientific knowledge by realigning the process <b>to not necessarily depend on the peer-review regimes</b></p> <p><b>Openness, inclusiveness</b> → <b>enable smaller institutions to participate</b> in the scholarly discourse</p> <p><b>Openness</b> → enables to <b>bypass the peer-review regime/bias</b></p>

### 8.5 *Participant A19*

Scientist A19 is an Assistant Professor in the Department of Physics, at a medium research university in the United States, with research interests in astrophysics and stellar astronomy. A19 received her Ph.D. in 2004, and her name is associated with about 40 deposits in arXiv, either as an author or as a co-author. She uses arXiv and ADS in her daily scholarly work, as well as raw data repositories such Chandra.

#### *Impact on scholarly process (A19)*

“And if I had waited – so I submitted the paper to astro-ph at the same time I submitted it to the Journal, and if I had waited for that manuscript to come out in the Journal, it would have been six months later, the conference would have been over and I never would have had the invitation.” (A19)

Scientist A19 perceives that arXiv had strongly performed on her career as a physicist, boosting her visibility within the scholarly community. She had submitted specific pre-print to arXiv and as a result she had been invited to a conference to present her findings. Timing was right for A19, as the international conference would not have invited her if she had waited for her paper to appear in the commercial journals. In this case, the conjunction of schedule timing and the availability of arXiv for depositing pre-prints, has realigned the career path of A19 to be recognized by the international community, reflected in her article, that has entered the disciplinary knowledge ecosystem much earlier—here we see the performative capability of this particular article potentially reconfiguring the knowledge network of other individual scientists.

“... open access citations are very critical, because the work gets posted to astro-ph so much faster than it would come out in the journal. And so it, I'm able to access critical information, build my conclusions around it, and submit a new

manuscript in half the time that it would have taken to wait for that article, article to come out in print.

And actually, I probably would still be – like okay, this past December, I wrote an article oh, with my graduate student and we had, we had a specific goal that we were studying regarding these stars and we would have submitted the article in December anyway.” (A19)

“And it just so happens that this paper came out, let's see what date it came out. It came out November 15 on astro-ph and this did not come out in any journal yet. I mean its February and I still haven't seen it out in a journal.” (A19)

As it has been expressed by other participants, the main performative aspect of the open access as manifested in arXiv and ADS is with respect to time and speeding up the research process, in addition to enabling open citation linking and aggregation of sources in one place. This capability has been utilized by the scientists by realigning their processes too easily and quickly search, discover and access materials for their work. A19 places arXiv at the center of knowledge exchange and sharing, both for accessing knowledge for local use and for the ability to quickly contribute research findings to the disciplinary knowledge network.

“And it said, ‘Here is the impact of those uncertainties.’ As that paper was coming out, I was collecting data to study these system geometries of that star and so when that paper was released on astro-ph it helped motivate my own direction to approach that research. And this happened just this past fall so I read the paper on astro-ph in November or December and very soon after I was able to submit a paper that addresses those concerns. So without having to wait for the work to come out in the Journal I have been able to speed up my research. ” (A19)

Again, a recurring theme is speeding up research, enabling faster production of an article and its wider and quicker distribution, and all of these by using scholarly knowledge artifact that may not have passed the peer-review process yet. The performative capability of arXiv to empower scientists to sidestep the peer-review process goes beyond the reduction of the publication gap. It also positions arXiv as an



actor capable of speeding up research, a role that has traditionally been associated with commercial publishers or other providers of value added services such as citation indices.

“But we were able to discuss the applications of our work with respect to this paper in December when we submitted our article. And had we not had open access we would have, we would have still submitted our paper but we would not have been able to investigate the implications to a broader audience.” (A19)

A19 perceived that the open access makes a tremendous difference for her research where there seems that there is no excuse for not knowing the latest relevant research. This is a rather strong reconfiguration of the perceptions amongst the astronomers and astrophysicists with respect to the level and comprehensiveness of the literature review to be performed for any given research paper. The expectation of always having comprehensive and complete ability to see the latest research was not the case before arXiv was available. Thus, arXiv has enabled the emergence of an expectation that performs upon scientists’ production processes—forcing upon them to do a thorough search of relevant literature as it is expected of them to do so.

“Right. Astro-ph, that’s true, but I think there was a study published on astro-ph that looked at the citation rates of articles that first appeared on astro-ph in later refereed journals ... Versus articles that never appear on astro-ph. And so the citation rates of articles that first appeared on astro-ph is much higher.” (A19)

Here A19 references evidence that commercial journals’ articles that initially appear as pre-prints in arXiv have higher citation impact. The performative capability of this evidence further encourages scientists to continue the process of using materials from arXiv and contribute therein. The performative capability of arXiv in this sense has reconfigured the scholarly production process to accept arXiv as an integral part of the knowledge production context in physics.

“Well I know that it's helped my career. I don't know how specifically it's changed scholarly communication but I do know that posting on astro-ph increases the visibility of papers.... Because, the audience reading the papers is much wider.” (A19)

Further, A19 links the wider dissemination of the pre-prints and post-prints via arXiv to the increased visibility of papers. This suggests that the quicker and wider dissemination as properties of arXiv translate themselves into the increased visibility for a specific article. As it will be shown later, the translation process of these properties is carried through the technological level of the access tools that would include the abstract and immediate full text search without the interference and filtering by the commercial journals or commercial services. The most prominent access tool in this case is ADS that indexes abstracts and provides full text search for arXiv materials.

*Impact on scholarly output (A19)*

“I guess my role in using open access is to both submit papers and to read other people's submissions. And when I read other people's submissions, it often motivates my own work. So a recent example of that was a paper about a particular star that I've been interested in for some years. And that paper looked at some uncertainties about the geometry of that star's orbit... so when that paper was released on astro-ph it helped motivate my own direction to approach that research.” (A19)

A19 perceives that the availability of different sources in arXiv as well as the availability of pre-prints and post-prints trigger new ideas and motivate her own scholarly work. In this specific case, arXiv submissions have been able to refresh or reinstate A19's interest to work on a problem she had been curious for some time, by relating to other scientist interested about the same or very related problems.

“By a large amount and I think the citation rates of your papers go up [by posting to astro-ph]. There have been studies actually post[ed] on astro-ph that show the citation rate goes up when you post astro-ph opposed to waiting for it to come out in the Journals.” (A19)

A19 perceives that the citation rates of articles deposited in arXiv increase, enabled by arXiv's ability to perform on the articles' discoverability. Thus, arXiv has removed the barriers from articles so they can be discovered by being included in an open information environment, by realigning the knowledge artifact topology (comprised of articles, journals, repositories, pre-prints, etc.) to be more inclusive.

“But I like the astro-ph because I – sometimes I don't necessarily know exactly what I am looking for. Sometimes I want to be motivated by a new random idea that I may not have thought of before. And that's why reading astro-ph is so important. Because people are very creative and they think of things that I never might have thought of. And so if I see a paper on astro-ph I can – it can trigger my own new ideas and I would say that my research is influenced by both. ” (A19)

A19 perceives the importance of astro-ph with respect to enabling A19 to read random articles, or set of articles that either alone or in groups can trigger new thoughts and ideas. Such triggering of ideas certainly happens by reading articles in commercial journals. The value and role of arXiv as a trigger of new ideas emerges from its ability to aggregate multitude of sources and enable an easy way to search the networked knowledge artifacts.

“And sometimes it's the data through these open access repositories actually can motivate further study. For example, a few years ago I did a study of some runaway O stars, some runaway massive stars. And I found one system that was moving very, very, fast through space like it could be a runaway star. But that star was classified as a member of an association of stars. So if it's a member of an association then it's not running away from the association. So that was kind of contradictory. So, I used the open access to find out more about the other known members of the association [of stars]. ” (A19)

In addition to the availability of pre-prints and post-prints that motivate new research, A19 perceives that the availability of astronomical open access raw data repositories can also trigger new thoughts and ideas. Astronomy and astrophysics, as

theoretical and practical disciplines base their research on raw data such as telescope observations, radio observations, astronomical x-ray data, etc. The availability of these raw astronomical data via open access enables them to be linked and associated at technical level with various knowledge artifacts—an example of which is ADS where each entry for a specific article also links to astronomical object repositories such as SINBAD (hosted at Harvard University) and NED (hosted at California Institute of Technology), raw data repositories such as MAST (hosted at The Space Telescope Science Institute), in addition to its links to arXiv, references and citations. This reconfigured and constructed network topology has emerged as a result of the electronic network environment in conjunction with the ability for resources to be interlinked to each other due to being available as open resources. Thus, the properties of openness and integration, as well as the property to be interlinked at technical level, has realigned the disciplinary culture by bringing forth a new actors (i.e., raw data and astronomical object repositories) that have become part of the knowledge network topology.

“If I didn’t have access to open access repositories I would probably be spending a lot of time in the library, looking up published printed journals, going through very tedious catalogues, line by line looking for available data measurements. And the thought of that is so discouraging that I would probably give up before I even started. Because some of these catalog that have hundreds of thousands of stars and dozens of different entries on the star.

It's extremely tedious work, it's easy to make mistakes, and that doesn't – some of the work that is posted online its still can't have mistakes but it's less frequent. And you should still be cautious but there's a lot more human error introduced if every single human is going and doing the searching themselves. And it would severely limit the volume of data that I could analyze and the volume of various scientific conclusions that I could make, I think. ” (A19)

A19 very clearly emphasizes the role of open access vs. not having open access. A19 perceives that open access has enabled her to look at many more data catalogs, and it has reduced the time for searching, discovering and accessing of astronomical data

tremendously, thus enabling her to make more scientific findings and conclusions. Primarily enabled by technological capabilities, the open access electronic networked topology of raw data repositories and data catalogs enable scientists to analyze much larger amount of data in shorter time periods. Thus, in this case it is the openness and the integration properties of open access data repositories (in conjunction with the technological capabilities) that have realigned A19 personal knowledge production process, translating a large amount of data into a manageable workable set of resources so that they can be considered for use in a broader set of research problems.

A19 perceive that the capabilities that have been provided by the open access repositories of articles have enhanced the quality of research: “So I think it [open access], I think it's enhanced the quality of our research. It enabled us to target a wider audience of readers” (A19). Increasing citation rates, as was mentioned earlier, is quoted as one measure of enhanced quality. Thus, open access as a property of arXiv and ADS has the ability to realign scholar’s knowledge production process by making the articles that are deposited therein more visible and thus more easily can enter the knowledge networks of more researchers.

*Integration with scholarly context (A19)*

“Well I know that it's helped my career. I don't know how specifically it's changed scholarly communication but I do know that posting on astro-ph increases the visibility of papers.” (A19)

“I think that – my – I think my wow moment was when I submitted a paper a few years ago and a week later had the invitation to speak... that to me was just amazing. And I really saw the power of the open access to influence my career and influence my reputation in the international scholarly community. ” (A19)

As it has been already mentioned earlier, A19 perceives that the posting of pre-prints and articles to astro-ph increases the visibility of the papers, and in her specific case has significantly performed on her career path. A19 perceives very strongly that her posting to arXiv jumpstarted her career in a way that would not have been otherwise possible. arXiv provided the venue through which A19 distributed her article and as a result was invited to speak at an important international conference. From this perspective, arXiv acted as the translation actor through which A19 career was realigned in conjunction with her individual knowledge production context.

“And I would say it's an ingrained part of the astronomy culture to post to and read astro-ph every day, but I don't think the same culture exists in physics.” (A19)

“Yeah, it's just a famous part of our culture and I'm training my graduate students here to consider astro-ph as one of the primary tools that you need to do your research.” (A19)

According to A19, there is also a strong perception that reading and posting to astro-ph is part of the disciplinary culture of astronomers and astrophysicists, and that the new entrants need to be initiated into the culture. Thus, by becoming an important and relevant actor at personal and local level, due to the social, technological and organizational aspect of interconnectedness of the actors that comprise the disciplinary scholarly communication production process as well as the disciplinary knowledge ecosystem, astro-ph has entered and has become part of the disciplinary knowledge production context: “They know that they can find 99 percent of the articles on astro-ph” (A19).

A19 shared a story that is unique, where astro-ph, due to its integration with the everyday life of astronomers in terms of research, has enabled the emergence of a new research actor, the “morning coffee”:

“I know the astronomy department at Ohio State University they are actually famous for holding daily morning coffee's where for a half hour everyday they discuss the top postings on astro-ph. ” (A19)

“And every time someone at Ohio State's astronomy department submits a paper in the acknowledgements they almost always thank the morning coffee discussion for having spurn the idea that spun that paper and the morning coffee discussion centers around astro-ph. So they are actively using the open access resources, discussing it, and following up on it with new research. ” (A19)

The astronomers at Ohio State University review on regular basis every morning the latest postings to astro-ph. As a result of these meetings, new ideas emerge and new articles are published. This new phenomenon is so ingrained in the daily research life of astronomers at Ohio State Universities, that the event—the morning coffee (“M. Coffee”)—has been listed as “author” in a number of papers that have emerged as a result of these daily meetings. The research process has been realigned in such a way to give rise to a new type of authorship: the morning coffee. M. Coffee is an example of a collaborative use of arXiv—a mini conference in sense that department members get together every morning, scan the new items in astro-ph and provide their input to fellow researchers, but sometimes unique ideas emerge that trigger new research.

*Democratization of the scholarly discourse (A19)*

“Well I think one of the negative aspects is that without open access you are relying on your school's library to have a variety of collections available for you to peruse. And here at [A19's institution] we don't have a strong astrophysics department, so our library is very weak on resources. So and even our electronic journal subscriptions are sometimes not as extensive as they should be.” (A19)

“So access to astro-ph probably doubles or triples the number of articles that I can have access to just here at [A19's institution] because our library doesn't subscribe to so many journals. So that's been very helpful like for example, Nature articles before 1992 we don't have access to here at [A19's institution], I don't know why but if that Nature article was posted on astro-ph then I can read it through the astro-ph server or the journal PASP, The Publications of the

Astronomical Society of the Pacific. We don't subscribe to PASP so if it gets posted to astro-ph then I can read it. And that's a big help sometimes.” (A19)

A19 perceives that even without a strong astrophysics department at her institution she is able to participate successfully in the disciplinary discourse. This would suggest that arXiv is perceived as a local scholarly resource by the department, without having to rely solely on the resources of the institution to provide access to the scholarly artifacts necessary for the research of its faculty. Thus, the open access repository has opened up the scholarly production network topology to allow for invisible colleges of scientists to enter the network, irrespective of their institution's buying power of commercial journals.

*Participant A19 summary in matrix form*

As it has been already described, Table 11 is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A19, in relation to the four themes.

Table 11: Summary of OA properties and their relation to the four themes as perceived by A19

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A19</b>	<p><b>Timing, Visibility</b> (of pre-print), <b>open dissemination</b> → bust if scholarly career, invitation to international conference</p> <p><b>Time, space</b> (early access; <b>quick</b> and <b>wider</b> dissemination), <b>integration, linking</b> (arXiv, ADS) → <b>increased visibility</b>,</p>	<p><b>Openness, integration, inclusion of pre-prints</b> → trigger <b>new idea and motivate research</b></p> <p><b>Time</b> (quicker, earlier distribution and availability) → enables to publish earlier than what would have otherwise</p>	<p><b>Openness</b> (posting to astro-ph) → <b>increases article visibility</b></p> <p><b>Central role</b> → <b>expectations to post and read from astro-ph</b></p> <p><b>Openness, central role</b> → “<b>morning coffee</b>”</p>	<p><b>Openness, visibility, trust</b> → <b>realigned the value</b> of an small institutional department – enabled <b>inclusion of scientists</b> in the scholarly discourse</p>



	<p><b>discoverability and accessibility</b> of articles</p> <p><b>Openness, linking, early</b> availability → <b>accelerated research</b></p> <p><b>Complete, open</b> availability of latest research → no excuse for not knowing the latest research</p> <p><b>Openness</b> (arXiv) → citation impact increase</p>	<p>been possible</p> <p><b>Openness, integration, inclusiveness</b> → higher <b>citation impact</b> (career steps realigned; linking and technological levels realignment possible)</p> <p><b>Open availability of raw data, linked, integrated</b> → more <b>relevant and complete</b> research</p> <p><b>Time</b> (quicker finding, discovering, accessing articles) → enabled more time to be spend on research and writing, able to analyze larger amounts of data</p> <p><b>Openness</b> (article, raw data) → <b>enhanced quality of articles</b> that are produced</p> <p><b>Openness, linking, early</b> availability → <b>accelerated research</b> (pre-prints not gone through the peer-review process enter the knowledge ecosystem)</p>	<p>phenomenon (new actor/structure has emerged; new author has emerged; scientists' daily routine has been realigned)</p> <p><b>Openness, central role, trust</b> → Critical venue for <b>career boost</b></p>	
--	--	---	--	--

## 8.6 Participant A20

Scientist A20 is Tenured Astronomer at a large observatory in the United States. A20 received her Ph.D. in 1996, and her name is associated with about 30 deposits in arXiv, as either an author or a co-author. She uses arXiv and ADS in her daily scholarly

work, as well as raw data repositories such as Chandra, raw data access tools such as SINBAD, and various databases at her observatory. In addition to research, she also manages a software group that builds software for managing telescopes.

*Impact on scholarly process (A20)*

“Sometimes it [use open access] is the most critical thing that I do. Like I have collaborations, occasionally, in which I have never met my collaborators. And the only way that we communicate is by email and sending papers back and forth. I put things up on the Web for them and you know – I mean, of course, occasionally, we’ll have to have phone calls but mostly it’s online only communication.” (A20)

When asked during the interview about the role of open access in scholarly communication, A20 perceives a relationship between the availability of open access and the ability to collaborate with peers around the world on various projects and articles. A20 perceives open access as a set of communication tools enabling collaborative research processes, more specifically as an enabler of collaboration with scientists with whom collaboration would have been very hard if the traditional communication tools (such as fax, phone and postal mail) had to be used.

“Well, when I first started working in the ‘70s there really wasn’t very much open access. In fact, computers were just coming out. And we did software development on punch cards. You know. That was it. And email was really quite rare. I remember when I first got my first machine I had to start learning how to do email. And you know things have changed a lot to the point right now we rely so heavily on it I couldn’t do my research without open access.” (A20)

Here A20 provides a historical account of the cultural configuration of astrophysics and astronomy with respect to the type of information technology used by the scientists, reflecting on the time when there was no electronic e-mail and software was developed on punch cards, and there was no electronic networked environment even for e-mail communication. For A20 there seems to be a dimension on which availability of

electronic networked environment and open access is associated with collaborative model of work practices. Perhaps, because open access has been ingrained so much in the personal culture of astronomers and astrophysics and their respective disciplines, and because arXiv has been used by scientists since 1991, they perceive open access as “natural” outcome of the innovations in the electronic networked communication environment.

“But at the time, I would do things. I would make a lot of phone calls. You know I would plot things on graph paper and have typists type up the report. Those reports generally didn’t go out very many places. I relied heavily on paper copies of journals in libraries. And I would go sit in libraries for long periods of time with books spread out all over the place. You know that’s what those big tables are for, right, so you can look in different books. And relating what’s going on in each one. And then I would make photocopies of things so that I would work not in the library. And I would have much, much more limited access.” (A20)

In order to explain the value of open access, A20 related to specifics of her research process based on paper and typing machines. Her personal knowledge production process had been restricted in two ways. One, the mechanics of searching and finding articles across different resources and then arranging them into an article for publication were very slow and tedious. Time as a property of the research process was mostly spent in searching, finding and cross-linking vs. argument building and interpreting. Second, the dissemination of the results was expensive and limited. The availability of arXiv has addressed both of these issues for A20. It has performed on the property of time by “shrinking” the time it takes to research any particular problem through literature review; it has “shrunk” the distance and space across resources by making distant and global resources very local and immediate; it has increased the number of resources and articles she can consult; and it has also alleviated A20’s activity in relating articles and

raw data sources to each other to extract meaningful and probable scholarly relationships by automating citation and bibliographic linking.

“And there were also fewer collaborations, the number of authors on each particular paper was smaller because you had to work with either mail and phone. You didn’t have email. And I couldn’t establish international projects, for example, very easily, whereas now, at least half of my collaborators are international.” (A20)

A20 also perceives that open access has opened up a number of collaboration possibilities, thus now there are articles that contain a larger number of collaborators (co-authors) compared to the time when open access in physics was conducted with paper based pre-prints. Thus, open access has realigned the authorship process to enable more researchers from great distances to engage with each other’s work in collaborative fashion.

“Yeah. Usually if I see something in a closed access I’ll read the title, I’ll read the abstract because usually the abstract is open access. And then I’ll think very, very hard whether or not I absolutely need it. Whereas, if it’s open access I’ll just download it and start skimming it. And when I do the first skim, you know, I’ll read like the intro a little bit, and then I’ll look at the figures, then I’ll look at the figure caption, and then, only then, if it still looks interesting, then I’ll go back and start reading the paper. Otherwise, I drop it at that point.” (A20)

“Whereas, if it’s not open access, very often, I won’t use that download. You know and I find a lot of things that are quite interesting, you know, that just keep my interest or something. And if it’s easy to download then it broadens my research avenue. But if it’s, yes, if there’s boundaries to me getting the paper I tend not to do that unless I know I absolutely have to.” (A20)

A20 experiences the performative aspects of open access at individual level in her decision making process on whether to consider an article for use or not. The availability of articles and pre-prints in open access has enabled A20 to quickly review an article beyond the title and the abstract, so it can be considered for more detailed reading. Because most open access repositories either carry the article in full text or have a link

to the full text location, a researcher can make a local copy of the article by downloading it. The local availability through downloading the article performs on A20 decision process by becoming part of her localized knowledge network. Instead, because closed access resources give you only article's title and abstract and make it harder for the article to become part of her local knowledge network, for A20 considers close systems such as commercial journals as barriers for her further consideration of the articles that are not open access. Thus, the full text availability of articles in arXiv has direct impact on increasing their visibility.

“Yes. And then I let the paper sit for sometimes months on end until I could figure out where I could get the information. A lot of times, when you get observations, the observations don't give you what you thought they would. You know you write this proposal, and you say you're going to get these observations, and you're going to answer this question. Then you find out the question isn't even addressed by the observations. And then you're left with month's worth of work that you're trying to figure out, what on earth does it mean. And you have no clue. Right? And you do searches and things and you can't figure out what's happening. ” (A20)

“And then, what I do is I fall back – I put it aside because usually I have four or five projects going at one time on all different levels. And often what I'll do is I'll start presenting posters at the two main meetings. So I'll start presenting posters and then I'll start talking to people. I'll try to talk to extra galactic and galactic type astronomers, theoreticians, and observers, show them what I'm doing, some ideas on it, and a lot of times it's through that kind of informal discussion that they'll give me, you know, the beginning of a new search or a new idea for interpretation. ” (A20)

A20 also perceives that the collaboration practices enabled by open access are instigators of new ideas, thoughts and articles, by enabling the scientists to work on multitude of problems concurrently by sharing preliminary results with colleagues and collaborators through conferences, presentations as well as informal communication. Thus, A20 has benefited from her drawn out knowledge production processes around

specific challenging research problems over a longer period of time, by being able to consult formal and informal scholarly channel due to their open nature.

*Impact on scholarly output (A20)*

“So at the beginning I usually have some sort of an idea either based on observation, or a talk that I heard, or something I read in another paper associated with some source or some problem in astrophysics. Then I do research to figure out whether that idea is reasonable. So mostly that’s with ADS looking at what other people have done, trying to figure out what’s going on in the particular region. Whether anybody can serve this source in this particular way.” (A20)

A20 includes ADS in her search process to find out related article around and idea she is interested in. ADS is thus able to perform on the types of ideas and research problem that she will work on, by using ADS to scope an idea by looking at what other scientists might be working on. Using the language of ANT, ADS is the conduit through which researchers can observe ideas in making and helps researcher impact each other with ideas that might not be published even as pre-prints, acting as an invisible college, where individuals researchers inform each other via individual “work-in-progress” networks.

“Yeah. If you want to know the history, the background, get the big picture, you know, what’s important about the observation that you’re doing, for example. How does your work fit into the larger quest for knowledge about the universe? You have to use things like ADS and you know the journal articles.” (A20)

For A20 ADS plays a crucial role in the research process as it enables her to build a complete picture around a specific astronomical image. Two critical properties of arXiv and ADS emerge here: a) collaborative, integrated and comprehensive resources, and b) the newness of the research.

A20 perceives that pre-prints are very critical for the knowledge network, it enables scientists to find and access the most recent and relevant articles for their own research: “Preprints are critical. Especially, you know when I do my papers I want to cite the most recent and most relevant things” (A20). Here we see the emergence of two open access properties in A20’s research process: a) finding recent research, and b) finding relevant research. Both of these open access properties have the ability to perform on the types and scope of research problem A20 decide to work on.

*Integration with scholarly context (A20)*

As it has been already mentioned, A20 uses arXiv and ADS at the beginning stages of article writing as well as in the later stages to contextualize the findings. “Yes. And I hate that” is how A20 shows her frustration when she needs to look for article in a commercial journal (especially when she is traveling) because its pre-print or post-print is not available via open access. The availability and use of arXiv and ADS are so ingrained in the scholarly production process, they are trusted to reflect the quality of the commercial journals via the pre-prints and post-prints, that any need and attempt to reach to commercial journals directly to access an article is not a welcomed step and it is taken only if necessary.

*Participant A20 summary in matrix form*

As it has been already described, Table 12 is a brief summary contextualizing the properties of arXiv, ADS and related tool and resources as experienced and perceived by A20, in relation to the four themes.

Table 12: Summary of OA properties and their relation to the four themes as perceived by A20

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>A20</b>	<p><b>Openness → collaboration, new ideas</b> and thoughts triggered</p> <p><b>Electronic networked environment → open access</b></p> <p><b>Time, openness → speedup the search process</b> for article</p> <p><b>Openness, time (quicker, wider dissemination)</b></p> <p><b>Openness → collaboration</b> (articles with more authors than before; artifact ecosystem realigned; article as artifact realigned)</p> <p><b>Openness → increases article visibility</b></p>	<p><b>Linked, integrated →</b> enables scientists to <b>find related articles; comprehensive view of a research problem</b></p> <p><b>Openness, time (quicker, wider dissemination)</b></p>	<p><b>Completeness, central role in the discipline, trust →</b> arXiv, ADS central role in scholar's research process (commercial journals not visited as a result; disciplinary knowledge production context realigned)</p>	



## **Chapter 9. Analysis and interpretation of individual PhilSci participants**

The process followed in this chapter is equivalent to the one followed in Chapter 8 that described and interpreted arXiv group of participants with respect to the four themes.

In this chapter, the participants are presented in the following order: Ph10, Ph24, Ph26, Ph29, and Ph32. As with the arXiv participants, the non-sequential order of participant codes has been preserved for practical purposes considering the grounded theory coding was started before the recruitment process was completed. Also, as the analysis of each individual case progressed, these codes made it easier to associate the quotes and experiences with individual participants.

### *9.1 Participant Ph10*

Researcher Ph10 is a Ph.D. student in the Department of History and Philosophy of Science at a large research university in the United States. Her primary interests are in philosophy of physics and early modern philosophy.

#### *Impact on scholarly process (Ph10)*

“It’s [philosophy of science] a very technical literature, and it’s [PhilSci] the best resource that there is for that sort of technical literature, especially the cutting edge research because a lot of times, you find manuscripts there that haven’t been published yet.” (Ph10)

Ph10 perceives that PhilSci is valuable because it contains the latest and up-to-date research for philosophers of science: “The sort of cutting edge access is really nice” (Ph10); especially research that has not been published yet: “I like their papers there that

I can't find elsewhere" (Ph10). PhilSci's property of openness is related to the ability to disseminate the latest research before it is published in commercial journals. Thus, Ph10's individual knowledge production process is being realigned to use PhilSci in order to read and use the latest research findings much earlier.

In a sense, Ph10 perceives that PhilSci may have the ability to realign the scholarly production process by reducing the publication gap and by displacing some of the traditional roles journals have acquired and exercised over the years: "You don't have to wait so long for a journal to approve papers before you get to see what people are working on" (Ph10). The performative ability of PhilSci to enable view into "what people are working on" by reducing the publication gap is perceived as one of the major benefits provided by PhilSci. There is also recognition that PhilSci is opening the boundaries to research by enabling sharing of "cutting edge research" by bypassing the gate-keeping function of the commercial journals.

In response to an interview question whether Ph10 will prefer to use articles found via open access or via closed access journals, Ph10 perceives that her search process and writing stages do not change. Ph10 will use an article independently of the source it was found and accessed from: "I mean I won't hesitate to use one over the other. They're both equal" (Ph10). Thus, it would seem that Ph10 assesses the value of the article to determine whether it is worth using it for the production of specific knowledge artifact. The self-assessment of articles for scholarly use might be a trait of experienced scholars who are very familiar with the discipline and its epistemic culture, or perhaps a scholar trusts PhilSci to contain quality scholarly material. In this case, Ph10 is a graduate student, who values PhilSci greatly for its ability to provide quick access to the latest

research finding. With regards to using pre-prints from PhilSci, she expects the pre-prints and manuscripts “... to be finished, published works...” (Ph10) by the time they appear in journals. Thus, although for her work in progress she “won’t hesitate” to use materials from PhilSci, she still associated the quality of articles as emerging from the peer-review process by expecting the materials she is using to appearing in a journal.

*Impact on scholarly output (Ph10)*

In a response to a prompt during the interview to further elaborate on the role of open access in her research process she stated that although it is useful it is not indispensable: “No, it’s helped a lot and it’s made the papers better, but there’s never been a context where I just couldn’t write the paper without open access.” (Ph10). While Ph10 perceives that there is value in having access to the latest research much earlier and before it appears in commercial journals, Ph10 also perceives that there has not been a personal situation or a context that an article could not have been written without open access. It appears that there are two distinct concepts being expressed here. One, at repository level there is recognition that there is value in PhilSci because of its ability to enable scholars to read the latest research much earlier. Second, at article level there is recognition about the value of having access to the latest knowledge and “see what people are working on.”

“Okay, I think, I mean my discipline, you know, I’m still kind of new here [graduate student in the department of history and philosophy of science], but as far as I can tell, there’s a lot of sharing of preprint and manuscript versions of papers among philosophers. And there are conference proceedings that get posted on the PhilSci archives all the time, which are unpublished versions of papers that are made public prior to conferences. And so a lot of times, they’ll, the fact that they’re up on the archives will allow people to read them beforehand and come in with more informed discussion points later, so it does move the field in that way.” (Ph10)

Ph10 perceive that many in the philosophy of science value PhilSci for the ability to share manuscripts, pre-prints and post-prints, as well as share conference proceedings before a conference takes place. This enables scholars to come more prepared for the conferences. Thus, it emerged that Ph10 perceives a relationship between the ability to read and share pre-prints, post-prints and conference proceedings (as enablers of more “informed discussion”) and the enhanced progress made in the discipline.

*Integration with scholarly context (Ph10)*

Ph10 also expressed some notions that PhilSci is a continuation of an existing culture within the philosophy of science discipline to share pre-prints and manuscripts: “there’s a lot of sharing of preprint and manuscript versions of papers among philosophers” (Ph10). This would suggest that there is a point of integration of PhilSci into the existing scholarly exchange process where PhilSci has enabled the scholars to exchange early versions of their preprints and manuscript electronically. Although there is some integration with the scholarly context in the discipline and PhilSci is valued for its ability to provide a view into the latest and cutting edge research for individual scholars, the relationship of PhilSci with the disciplinary scholarly production process seems to be rather weak, as Ph10 has stated before that she can write an article without depending on PhilSci.

“Well, with the PhilSci archive itself, I feel very confident in the papers that I find there, and I will use them. I’ll cite them in my own papers, and I don’t see a problem with having them stand as they are in the PhilSci archive and expect them to be finished, published works by then. (Ph10)

As it was mentioned earlier, Ph10 states that she would assess the value of the papers she finds in PhilSci and would decide whether they are worth for referencing in her own

work, with the expectation that they would be published soon in commercial journals. Perhaps this is an explanation about why PhilSci is not perceived as important for writing articles. It is viewed as a transitory place for articles in progress on the way to be published in journals.

“If there’s a journal that they’re coming out in, I will designate the journal that they’re coming out in or I’ll just say it’s a manuscript that I found in the archives, but I mean you do, you do the same thing that you would with a lot of unpublished documents or documents that are forthcoming. It’s not like the PhilSci archive is the final resting place of the papers. Does that make sense? (Ph10)

Ph10 further clarifies that she does not view PhilSci as a publishing place, rather as a transitory place for articles in their lifecycle to move towards peer-review and then publication in commercial journal. Thus, PhilSci enhances scholars’ information practices but its perception of value with respect to the disciplinary knowledge production context is perceived as limited by Ph10—viewed a repository for conference papers and as a collaboratory.

#### *Participant Ph10 summary in matrix form*

As it has been already described, Table 13 is a brief summary contextualizing the properties of PhilSci and the related tools and resources as experienced and perceived by Ph10, in relation to the four themes.

Table 13: Summary of OA properties and their relation to the four themes as perceived by Ph10

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>Ph10</b>	<b>Openness, access → Latest, access up-to-date research</b> (some of which is unpublished )	<b>Access, time → quicker, faster access to latest research</b>	<b>Continuation of pre-print culture → old transposed into the new</b>	

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
	<p><b>Openness, sharing</b> → displacing commercial journal's traditional role</p> <p><b>Time</b> → <b>quicker, faster access to latest research</b></p> <p><b>Openness, early access</b> → <b>share conference proceedings and pre-prints</b> before a conference takes place (conference participation process realigned)</p>	<p>(BUT, the article could have been written even without open access)</p> <p><b>Openness, early access</b> → <b>share conference proceedings and pre-prints</b> before a conference takes place</p> <p><b>Openness, early access</b> → source of <b>new ideas</b></p> <p><b>Openness, access</b> → <b>Latest</b>, access up-to-date research (some of which is unpublished )</p>	<p>(a weak integration with scholarly process since articles can be written without using PhilSci)</p> <p><b>Place, collaboratory</b> → <b>transitory place for article on their way to peer-review and publication</b></p>	

## 9.2 Participant Ph24

Scholar Ph24 is an Assistant Professor in the Department of Philosophy at a large research university in the United States. He primarily works in the field of philosophy of science and bioethics. Ph24 received his Ph.D. in 2007, and his name is associated with about 10 deposits in PhilSci, either as an author or co-author.

### *Impact on scholarly process (Ph24)*

“There is often a very long lag between a journal accepting a paper for publication, and actually seeing it in print. Now, that’s starting to be dealt with by the journals themselves by posting things online first. So for instance, I had an article recently accepted in the Journal Biology and Philosophy, and even though that isn’t out in print yet, it is officially published and available online at the journals website.” (Ph24)

“But again, kind of two things – there is still even a lag between acceptance and seeing it online. And secondly, that isn’t available across the board in all journals. So the advantage of something like PhilSci archive is, once you know

that something's been accepted for publication, you put the penultimate draft up at the website, and basically it's a forum to get your ideas out there until it actually sees print." (Ph24)

As few other participants have already mentioned, scholar Ph24 also observes that some journals are responding to the open access phenomenon by distributing electronic versions of the accepted articles via their websites much earlier before the printed version of the journals come out. Nevertheless, Ph24 perceives that even so there is still a large lag between the time a paper is accepted and published by the journal. Therefore, PhilSci emerges as an actor with performative ability to realign the disciplinary knowledge ecosystem by injecting the newest research much earlier into the scholarly discourse. This has the potential to shift the current set of research challenges to be addressed earlier than what would have been possible otherwise.

When prompted about any other values that PhilSci may provide for his scholarly work, Ph24 emphasized strongly that PhilSci enhances the scholarly publishing process by enabling wider (space property) and quicker and faster (time property) distribution of scholarly materials: "No, I would say those are the two main ones [i.e., wider distribution of material, and early access to new knowledge before it shows up in journals]. Nothing else comes to mind." (Ph24)

"They certainly – I mean, as with anything on the internet, it's just a venue available to get information out there. And I've been contacted on multiple occasions, where people have encountered my work, or I will contact people because I've encountered their work because of browsing the PhilSci archive. And so I mean, in an academic environment, where you are judged based on your research activities, there's a real premium placed on getting your ideas out there as widely as possible. And so the fact that this is available really facilitates that endeavor that I think all academics have." (Ph24)

Even though Ph24 perceives PhilSci as a an environment that enables and instigates collaboration with the ability to connect scholars that share the same research interests

amongst themselves, a real emphasis is placed on the ability to disseminate own research results as widely as possible. Thus, there is a suggested relationship that wider distribution enhances scholars' research process and academic standing.

“And it’s really just a matter of looking through different resources that are out there that are designed to connect what you’re working on to what other people have worked on. PhilSci archive is one example of that, but there are other things.” (Ph24)

Ph24 perceives PhilSci as an actor that enables scholars to link their research with what other scholars have worked on. Thus, a property of connectiveness emerges for PhilSci by which other articles are performing on scholars' knowledge production process by introducing articles into scholars' knowledge network and making them visible for local use—some of which is translated into articles by the scholar.

“People will also then perhaps, use conferences as a venue to run ideas off, and then get feedback, and then submit that article elsewhere. But again, that may or may not get accepted somewhere. So the conferences option essentially let’s you get anything that you’ve presented at a conference, available for view, right. And so again, it’s a nice, handy access point to examine research that may, in fact, not even make it to publication at all.” (Ph24)

“Well, another kind of nice feature is it offers a venue to put conference papers that perhaps, might not see publication online. So for instance, oftentimes, folks will – well, a conference will be organized, and you can – everyone delivers a paper, and let’s say there’s 15 papers delivered at this small conference. At times, there will be efforts made to publish a volume of those works. I would say that happens less often than people hope it would. And also, it happens less than it does.” (Ph24)

An additional value of PhilSci is that it is perceived as an enabler of collaboration specifically for conference participants as well as a “publication” venue for conference papers. More than often ideas discussed at conferences (“run ideas off”) might not get published elsewhere, reducing their chance to enter scholarly discourse. Thus, PhilSci has the ability to perform on the knowledge networks of scholars and enable ideas



discussed at conferences to enter the disciplinary knowledge ecosystem, either through published articles or through the work in progress presented at conferences.

Ph24 does not determine the value of articles based on their source of access (commercial journals or open access for pre-prints, manuscripts and other work in progress that might be presented at conferences). Rather, the scholar would assess the value personally, often determining based on personal knowledge about the author: “No, I don’t. I mean, because I – when it gets to the point where I’m writing my paper, I see the authors works as equivalent regardless, of where they’ve come from” (Ph24). It appears that personal knowledge about the academic standing of the author, perhaps even an author’s expertise about specific research topics, are more important in assessing the value of a specific article, regardless of its publication status (still in pre-print or a conference presentation, or already published in a journal).

*Impact on scholarly output (Ph24)*

“So to list PhilSci archive as the publication place, I would find to be a bit odd, unless, they were citing something that is in the conferencing section, in which case – well, I would find even that peculiar because, again, I see PhilSci archive as offering a venue to make information available before it’s in print.” (Ph24)

Ph24 perceives PhilSci as a preparatory venue for materials that would come in print, instead of a publication place. PhilSci is perceived to have a limited value in a sense that it enables scholars to read the latest materials and also publish their own materials much earlier. Thus, PhilSci’s ability to perform on scholar’s individual knowledge production context is perceived to be limited to help the scholar to do their work quicker and faster (information processing enhancements), with little implication on the dynamics between scholar’s knowledge production process and scholar’s

knowledge network, and the disciplinary knowledge production context where Ph24's and other scholars' information practices are enacted. However, the process of earlier dissemination of ideas does not live in its own bubble; rather it is situated in a scholarly publishing context related to the discipline. The early publication of ideas would then trigger other scholars to use them earlier as well. Thus, from a perspective outside of Ph24, it would seem that PhilSci has a performative ability to realign the disciplinary knowledge network in a different way compared to the context where PhilSci is not available.

“And I don't think that the web master, that PhilSci archive take themselves to be in the publication business, or the peer review journal business. What they're doing is offering a venue for material that's passed the peer review journal, and to get out there. But they have no way to check and see, right? It's not like they get a submission, and then they go and contact a journal, and say, hey, so and so says this is coming out there, can you confirm that.” (Ph24)

Again, Ph24 perceives PhilSci as mostly a venue for disseminating materials that have passed the peer-review process. Nevertheless, the performative potentials of PhilSci have ripple effect as the early availability of articles (even in pre-print stage) means that they might enter the disciplinary knowledge ecosystem much earlier, including here conference papers and pre-prints even if they do not make it through the peer-review process.

#### *Integration with scholarly context (Ph24)*

As it has been explained earlier, the integration with scholarly context for Ph24 is perceived to be limited to the ability to enable articles to be distributed widely and read much earlier, with very little recognition about PhilSci potential to impact scholar's information practices beyond PhilSci's use as a tool for finding and accessing article. To

clarify this further, Ph24 did not perceive any value that may come out from being able to access knowledge artifact quicker (especially the latest research).

“I mean the, what the PhilSci archive offers is a searchable database of essentially, either things that are up and coming, or things that have already been out there in the philosophy of science. So it’s just a default that whenever I’m working on a new project, whether it’s in the early phases and I’m just kind of getting a sense of what’s out there, or in the late phases and I’m looking to actually kind of engage actual authors, I’ll go to the PhilSci archives, bring it up and usually just kind of post browse based on the subjects. And also, do searches based on different individuals that I know are working in the area to see if there’s anything out there that I’ve missed.” (Ph24)

Ph24 uses PhilSci in his search process to identify the latest ideas (“things that are up and coming”) related to his research interest and find what other scholars are working on (“what’s out there”) about related subjects. There is also a perception that PhilSci will help him to find anything that he might have missed so far through the different stages of article production, some of which may be found by locating the scholars with similar research interests. Thus, in addition to finding the current and latest articles, scholar Ph24 perceives PhilSci as a venue with the property of potential completeness related to his research interests. This seems a bit contradictory. Why would such a resource that is perceived as useful in finding potentially missing articles for scholar’s research problems not be more integrated in the discipline? It would appear that there is a disconnect between Ph24’s perception of PhilSci and how it helps him—it is easier to find and access quicker and faster the current and up and coming findings in philosophy of science. One way that this disconnect might be explained is that the perceived values and roles of PhilSci are isolated from the rest of the information work—an example of which would be that Ph24 does not use the knowledge obtained through PhilSci when writing articles. This does not seem to be the case as earlier Ph24 stated that there is “a

real premium placed on getting your ideas out there as widely as possible”, clearly positioning PhilSci as an actor in Ph24’s knowledge production process that is able to perform with its property of space enabling wider dissemination. This contradiction will be further addressed in later sections.

“There’s a journal called ISIS, which is a history of science journal publication. And every year it puts out what it calls the current bibliography, and it covers history of science, it covers philosophy of science, it covers social studies of science. And again, whereas PhilSci archive is kind of looking forward to these are the things that are coming out.” (Ph24)

In comparison to other scholarly resources and tools that help scholars to find what has been published up to a certain point in time, Ph24 perceives PhilSci as a venue to help scholars to look forward to what is coming in the journals. Thus, it would seem that PhilSci performs on Ph24’s knowledge network by complementing the materials that are available through the journals, with knowledge from scholarly materials (such as pre-prints, conference papers, and manuscripts) that are indicators of research problem that would be showing up in the journals in the future.

Further, in line with Ph24’s perception and lived experience accounts so far, Ph24 perceives that he has used PhilSci extensively, but could not recall a moment where PhilSci was very crucial in the knowledge production process: “I’ve certainly relied on it [PhilSci] extensively, but I can’t think of a time where it was like I had an aha moment because of something I’ve found there.” (Ph24).

“Hard to say. I think most scholars in the philosophy of science are aware of it. And so, it gets its use, and that’s a good thing. I don’t know what it would be like if suddenly, rather than 36 entries under genetics, it was 3,600. I think probably it would be a bit overwhelming. The nice thing about this is you have a sense of, if it’s here, it’s relevant, and it’s worth taking serious. ” (Ph24)

Ph24 perceives that philosophers of science are aware of PhilSci and that it is relevant. However, the value expressed seems to be limited to PhilSci ability to enable wider distribution of articles and pre-prints, where even a larger number of relevant articles are not necessarily perceived as being of increased value, rather that it might bring more challenges.

“Versus [PhilSci being compared] if you were to just Google the same keywords, get 1 million hits. And I find that distracting. So certainly, you wanna see something like that continue to grow as more and more research, and knowledge production are obtained, but you don’t wanna see it get out of hand at the same time.” (Ph24)

PhilSci is also perceived as a disciplinary search tool for philosophy of science used by scholars to find relevant materials for their research that provides more relevant search results than Google—PhilSci acting as a disciplinary filter (in the ecosystem of the vast information that can be found via Google) that makes visible knowledge artifacts for philosophers of science. There is also recognition that PhilSci should grow further in a balanced way to include only materials of interest to philosophers of science. Therefore, an additional property of PhilSci is its role as a search tool to help scholars further discover and access articles, by becoming part of scholars’ search process.

#### *Participant Ph24 summary in matrix form*

As it has been already described, Table 14 is a brief summary contextualizing the properties of PhilSci and the related tool and resources as experienced and perceived by Ph24, in relation to the four themes.

Table 14: Summary of OA properties and their relation to the four themes as perceived by Ph24

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>Ph24</b>	<p><b>Openness, early access</b> → pushed commercial publisher to <b>reduce the publication gap</b> or <b>distribute electronic</b> version of article <b>much earlier</b></p> <p><b>Time, space (faster, earlier, wider distribution of pre-prints)</b> → access to <b>latest research before they are published</b> by commercial journals</p> <p><b>Time, space (faster, wider)</b> → <b>enhanced</b> scholarly publishing process</p> <p><b>Space (wider dist.)</b> → enhances scholarly research process</p> <p><b>Openness</b> → <b>collaboratory and publishing venue</b> for conference papers</p> <p><b>Collaboratory</b> → <b>connect scholars with each other</b></p>	<p><b>Openness, linking</b> → <b>collaboration</b></p> <p><b>Openness, time</b> (early access) → <b>preparatory place</b> for articles that would be published in journals</p> <p><b>Openness, time</b> (earlier publication) → <b>faster distribution</b> of own research</p> <p><b>Time, space (faster, earlier, wider distribution of pre-prints)</b> → access to <b>latest research before they are published</b> by commercial journals</p> <p><b>Space (wider dist.)</b> → enhances scholarly research process and?</p> <p><b>Collaboratory</b> → <b>connect scholars with each other</b> (enabling knowledge that otherwise will not be published to enter the disciplinary knowledge ecosystem)</p>	<p><b>Time, space, openness</b> → <b>wider, quicker distribution</b> of latest research materials</p> <p><b>Time, space, openness</b> → enables to see <b>completeness of scholars research</b></p> <p><b>Limited role in the discipline and scholarly communication</b></p> <p><b>Openness, latest research</b> → enables <b>forward looking view</b> of what's coming up in the discipline (ability to realign disciplinary knowledge ecosystem)</p> <p><b>Openness, access</b> → yet <b>not very crucial</b> for knowledge production</p> <p>Scholarly community <b>aware of PhilSci</b>, but with <b>limited use for wider dissemination</b> of pre-prints and post-prints</p> <p><b>Openness, search tool</b> → enables scholars to <b>find relevant article</b> for their research</p>	

### 9.3 Participant Ph26

Scholar Ph26 is a Senior Lecturer in the Department of Philosophy, Logic and Scientific Method at a major research university in the United Kingdom. He primarily works in the philosophy of science. Ph26 received his Ph.D. in 2003, and his name is associated with about 10 deposits in PhilSci, either as an author or co-author.

#### *Impact on scholarly process (Ph26)*

“Because there is a very long backlog for many journals. I mean in some cases it takes up to two and a half to three years for them to actually appear in print from the moment it has been accepted. It’s a very long time, and just to make things accessible to the community in the meantime, I just put them in the archive. That’s really what I use archives for.” (Ph26)

“So I mean it is really just – well, how should I say – the way we use it is just to bridge the time between acceptance of a paper and it appearing in a journal. That is really what it’s for. So it doesn’t really have sort of a life of its own.” (Ph26)

For Ph26 the main use and value provided by PhilSci is that it enables scholars to access the latest research quickly by bypassing the “long backlog” between the time an article has been accepted for publication and its publication in the commercial journals. The statement that PhilSci “... does not have a life of its own” (Ph26) would suggest that Ph26 perceives PhilSci as complementary the existing scholarly publishing via commercial journals, providing the researchers with an alternative quicker access point to the knowledge artifacts.

“I mean two years is half a lifetime a graduate student, so it is very important to have things as soon as possible, so that is really a great help. I think the main advantage of these things is that they’re fast ... As opposed to the standards of channels of publications, which are extremely slow.” (Ph26)

In addition to reducing the time between article’s acceptance and its availability for use by the scholars, Ph26 perceives a value in the ability to get access to research articles

as soon as possible. Relating the value to the “... half lifetime of a graduate student...” doing graduate studies, it would suggest that the ability to access research findings much earlier than would otherwise be possible (due to the “long backlog”) has implications for the types of research and approach graduate students will undertake for their studies—enabling students to work on most current research problems.

“That is really the only use they [open access repositories] have systematically, just to make sure you don’t reinvent the wheel because as is often that case that their topic is sort of in the air. I mean ideas usually have a context; they don’t materialize out of nothing. And there’s always the worry that what you’re just trying to do has actually just been done by someone.” (Ph26)

Here Ph26 perceives an additional value stemming from the early availability of research finding: PhilSci is perceived as realigning scholars’ individual knowledge network by making visible to the scholar a more current and more complete set of knowledge artifacts (that have not appeared in commercial journals yet) enabling them to work on novel and unique problems and to complement each others’ work. The phrase “... don’t reinvent the wheel” is related to the inefficient approach to doing science and that the time can be used to research other research problems. Thus, the realignment of scholar’s individual knowledge network (by introducing materials from outside of the commercial journals network and by working of novel and unique problems) will have implications for the disciplinary context by reducing duplicate work and thus acting as an actor that performs on the disciplinary knowledge network enabling advancement and progress of the discipline.

“... I maintain a personal website, and so do most of my colleagues, and this becomes actually very important tool of information. So if I wanna know what Mr. so-and-so does, I don’t go to any archive or repository to find his name, I just go to his website.” (Ph26)



In addition to Ph26's perception that PhilSci complements the role of the commercial journals in his scholarly process by reducing the publication lag, his personal Home Page and those of his colleagues further complement the article search process in the instances when searching for materials by specific author. Thus, the Home Pages are emerging as a relevant actor in scholar's research process alongside PhilSci.

*Impact on scholarly output (Ph26)*

"I mean I only have access there to papers that only appear two years down the line, and if it's a topic I'm really working on, that's a huge difference." (Ph26)

"So go to the archives and check that there's nothing of that sort around [referring to similar problems currently being worked by]. I mean that is really the main use." (Ph26)

Even though Ph26 perceives a limited role for PhilSci in that enables earlier access to research findings, when the available articles are related to his own current projects, they make huge difference in Ph26's research process as a tool to make sure the scholar works on a unique problem and has access to the latest research. Here we see the open access repository's properties of time and as enabler of increased early visibility for materials deposited in PhilSci, performing on Ph26's individual knowledge networks by enabling selective process with respect to currently worked problems, and therefore providing Ph26 with an ecosystems of ideas (from the articles) that help him to select a research problem that is unique.

Ph26 relates the availability of articles earlier than what otherwise would be possible as being ahead of time with the latest research: "I mean the main use I see is that you just get things before they come out in journals. It is really being ahead of the time" (Ph26). Thus, as it may appear that the early access to research findings by itself is

isolated from other aspects of scholar's and discipline's scholarly production process, there is value in being ahead because those research findings can have impact on the type of research a particular scholar conducts. From the perspective of the totality of scholarship in the philosophy of science discipline, there are no advantages at personal level as all scholars that value PhilSci will know that it is the source for the latest research. However, considering that scholars and their individual knowledge production contexts are actors in the disciplinary knowledge production context, it seems that PhilSci would perform on scholars' individual knowledge networks—providing them with “fresh” materials that can trigger novel and unique ideas.

*Integration with scholarly context (Ph26)*

“It is moderately important. I mean it is still just a facilitator. I mean it has absolutely no prestige. I mean you can't say, “I wrote the paper and placed it in the archive,” and then – then expect the promotion committee is impressed with that.” (Ph26)

Ph26 does not perceive a direct value from PhilSci either at personal level or disciplinary level. Ph26 also does not perceive that there is even an invisible or less direct impact of PhilSci on scholarship, and it does not have the prestige of the peer-review process that is valued by promotion and tenure committees. As with the other PhilSci participants, there is recognition that PhilSci does provide a value with respect to enabling quicker, faster and earlier access to latest research and it also bridges the publication gap (acting as an alternative “publication” channel for pre-prints and post-prints), but this value is somewhat framed by researchers' perception of being “just a facilitator,” an information processing “machine” that is isolated from the rest of the scholarly knowledge production processes in the discipline.

“For me very little [role for open access] actually. Well, I may hack every now and then a keyword into the Pitt Archive, but that is only about the third or the fourth thing I do. I mean first you take Stanford Encyclopedia of Philosophy, which is online, and then you use the Routledge Encyclopedia, then you use Philosopher’s Index, which is very good.

That’s [Philosopher’s Index] not really an online repository, that’s an online database that has only abstracts of papers, not papers themselves. And then you sort of compile a reading list of things that you should look at, that you want to check out. So – well yeah, as I said, I mean I would look at the Pitt Archives every now and then, but it’s not really one of the major sources of information.” (Ph26)

That Ph26 does not value PhilSci as much as other resources that are available to him is evident in the above quote stating that PhilSci is not central to his research activities on daily basis. Thus, PhilSci does not seem to have entered Ph26’s personal knowledge production process as a critical actor, rather it is used for scanning of ideas and complementing the other information resources used by Ph26. Throughout the interview with Ph26 there is no indication that PhilSci is perceived any differently by the broader community of philosophers of science judging by the little role if any it plays in the promotion committees’ decision process. However, for the very limited time he does check PhilSci to ensure that he is not reinventing the wheel, PhilSci does have a role in the types of research problems that Ph26 decides to work on.

“That [‘being ahead of the time’] is their [open access repositories] main plus. I mean look, they can never compete with the proper journal archive, let alone an electronic library that has some commercial products in it in terms of quantity.” (Ph26)

Ph26’s perception that PhilSci cannot compete with commercial journals online portals is a reflection of the philosophers of science disciplinary culture, especially the publication function (that is still in the domain of commercial journals that provide quality control through the peer-review process), reflecting on the perceived perspective

about PhilSci as a tool for dissemination of pre-prints and manuscripts and not as a publication place of pre-prints and post-prints.

*Participant Ph26 summary in matrix form*

As it has been already described, Table 15 is a brief summary contextualizing the properties of PhilSci and the related tool and resources as experienced and perceived by Ph26, in relation to the four themes.

Table 15: Summary of OA properties and their relation to the four themes as perceived by Ph26

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
Ph26	<p><b>Time (faster publication) → shorten the time</b> between acceptance and availability to the community</p> <p><b>Time, access</b> → ability to <b>read latest research</b> right a way</p> <p><b>Time (early access to latest research) →</b> enables the scholars <b>not to reinvent the wheel</b></p> <p><b>Home Pages</b> as <b>critical actors</b> in scholars' production process</p>	<p><b>Time (early access to latest research) →</b> enables the scholars <b>not to reinvent the wheel</b> (disciplinary knowledge ecosystem realigned)</p> <p><b>Early access, openness</b> → scholars is <b>ahead of time</b></p>	<p><b>Isolated</b> ("no life on its own" → not part of the broader disciplinary context</p> <p><b>Limited value, information processing</b> value only</p> <p>Perceived <b>no direct value</b> at personal or disciplinary level</p> <p>PhilSci <b>not central</b> to his activities on daily basis</p>	

*9.4 Participant Ph29*

Scholar Ph29 is a philosopher with primary interest in philosophy of science. He is an Assistant Professor in the Department of Philosophy at a large research university in

Canada. Ph29 received his Ph.D. in 2006, and his name is associated with about 10 deposits in PhilSci, either as an author or co-author.

*Impact on scholarly process (Ph29)*

Ph29 perceives PhilSci as a venue that enables scholarly communication amongst the scholars, in addition to the archival function for published papers that have gone through the peer-review process: “It is one means of initiating intellectual exchanges of communication, apart from papers that have been formally published” (Ph29). The intellectual exchange that is framed by PhilSci as a venue for communication is different and complements the intellectual exchange that is framed by the formally published papers. Thus, PhilSci has the potential to perform on the disciplinary knowledge ecosystem by enabling ideas and thoughts not represented by formally published papers to enter the disciplinary knowledge ecosystem.

“So you wonder what role using open access material plays in my own research? ... Probably just a minor role.... The main value is, of course, that you can get hold of an article before they have been actually published.” (Ph29)

Ph29 perceives that materials found via open access play a minor role in his own research. However, Ph29 does attribute limited value to PhilSci for the ability to be able to read articles before they are actually published. The value of PhilSci is thus associated with its performative ability to make available articles in their pre-print or post-print format before they are actually published in the commercial journals.

“... most of the material I use is already published, and those articles that aren’t published yet, but are of importance to my work, are often items that some of my colleagues sent me, and I become aware of them in this way. So from that perspective, most of the open, purely open access repository is not essential to my research.” (Ph29)

Again, Ph29 perceives that although open access enables scholars to see materials before they are officially published, for Ph29 there is little value in that because he always uses already published materials for his research, and those that are not published yet have been sent to him by the authors directly. It appears that Ph29's invisible colleges provide him with the ability to see the materials related to his research before their publications. This would then suggest that PhilSci acts as an access tool alongside scholars' personal Home Pages, where scholars usually are aware of the main research areas in the discipline and the scholars working in the main research areas are well known. Thus, the ability of PhilSci to realign scholar's production process and scholar's knowledge network is reduced by the similar role that Home Pages play in scholars' research processes. In a sense, PhilSci and scholars' Home Pages complement each other, where Home Pages can be viewed as a distributed open access repository, where articles can be discovered in the context of the rest of author's works, whereas in PhilSci the "author" context needs to be built by scholar's queries.

"I hope that some people in the philosophy community, or philosophy of science community, can become aware of my articles before they are published, or they – even once they are published – see them in the open access repository first and, of course, they already see the abstract, and they can easily download it. My hope is that one of, kind off few would actually read my paper that way, but I have no statistics if this actually happens." (Ph29)

Here Ph29 further elaborates on the value of PhilSci. The ability for scholars to have access to articles and research before they are published actually is directional for Ph29. He perceives that as a result of his depositing his own work into PhilSci he may be able to realign the knowledge networks of other scholars. Himself however he does not see tremendous value from PhilSci, primarily because his invisible college is strong and many of his peers make their work available early through their Home Pages.

*Impact on scholarly output (Ph29)*

As it has already been mentioned, Ph29 perceives PhilSci as valuable because it enables scholars to read articles early before publication. As a property of PhilSci, enabling earlier publication of articles is not necessarily valued highly by Ph29 because he has the means to access the articles via his invisible college and the Home Pages of scholars. Yet, the early access property as assigned by Ph29 to his invisible college actually has the performative capability to impact his scholarly output by introducing into his knowledge network articles that otherwise might not be considered for the specific research problem. Thus, whether it is the invisible college, the personal Home Pages of scholars or PhilSci, these sources act as access tools to enable Ph29 to access and read the latest research much earlier than otherwise would be possible if scholars relied only on commercial journal publications.

“Probably rather [than] enabling the process, maybe, in that open access repository gives you access to an article before it’s published, so now and then, you take a look at articles that you happen to find, at least you read the abstract, maybe sometimes the article.” (Ph29)

Ph29 gets access to the latest articles relevant to his areas of research through his invisible college. This approach to finding the latest research might be restrictive from the perspective of serendipitous discovery by not being able to look at many different articles from different sources in one place. Thus, there seems to be a differentiation between the potential performative capabilities of PhilSci because of its ability to aggregate materials from multiple sources in one place, in comparison to Home Pages of scholars and the broader invisible colleges that enable targeted search for specific articles. All three sources for finding and discovering articles (PhilSci, Home Pages, and

invisible colleges) have implications for scholar's individual knowledge production contexts.

*Integration with scholarly context (Ph29)*

“I don't think there are any disadvantages, or I'm not aware of any. In my work, I primarily use open access repositories to – in addition of my Home Page – make my materials more widely available by posting a pre-print.” (Ph29)

The extent to which Ph29 perceives PhilSci as integrated with the scholarly context is its ability to make his materials widely more available so other scholars can read them and use them in their scholarly works. For Ph29 this property is complemented by his own personal Home Page where he also posts his own pre-prints and post-prints.

*Participant Ph29 summary in matrix format*

As it has been already described, Table 16 is a brief summary contextualizing the properties of PhilSci and the related tool and resources as experienced and perceived by Ph29, in relation to the four themes.

Table 16: Summary of OA properties and their relation to the four themes as perceived by Ph29

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>Ph29</b>	<p><b>Openness, communication</b> → enables <b>scholarly communication</b></p> <p><b>Openness, pre-prints availability</b> → enable the <b>exchange</b> of materials that are <b>not formally published</b></p> <p><b>Time (early access)</b> →</p>	<p><b>Early access</b> → enabling view of <b>latest research</b> (not so much value for this scholar since his invisible college is stronger; they perform on his knowledge network)</p> <p><b>Invisible college</b> → <b>restrictive to</b></p>	<p><b>Limited role</b></p> <p><b>Limited role even with the early availability of latest research</b> (usually authors send him the latest research directly)</p> <p><b>Home Pages</b> of authors more</p>	



	able to read <b>latest research</b> articles  <b>Time, space</b> (faster, wider <b>distribution</b> ) → enabling the authors to <b>disseminate</b> his work for other to read and influence their research	<b>serendipitous</b> discovery (my comments)  <b>Openness, pre-prints availability</b> → enable the <b>exchange</b> of materials that are <b>not formally published</b>	<b>integrated</b> with scholarly process	
--	--	---	--	--

### 9.5 Participant Ph32

Scholar Ph32 is a Professor in the Department of Philosophy at a medium master's level university in the United States. He primarily works in the field of history and philosophy of science. Ph32 received his Ph.D. in 1994, and his name is associated with about 10 deposits in PhilSci, either as an author or co-author.

#### *Impact on scholarly process (Ph32)*

“And also I would say the other one I use the most would be the PhilSci Archive because I do philosophy of science and a lot of the articles that will be coming out – or earlier versions of articles that’ll be coming out – are put up on that site, and those are easy to obtain right over – easily as well. And you can then kinda get an idea because they have sort of an update page of things that people are putting up there. You can kinda get an idea of what other people are working on, so it kind of gives you a feel as to what – what’s happening in the field as it were that I work in.” (Ph32)

Ph32 perceives that PhilSci enables access to current work in progress as it happens by other scholars in the discipline. PhilSci is then perceived as a conduit through which Ph32 stays up to date with the happenings in the discipline. Ph32 perceives a value in the ability to access research findings right after they are completed while otherwise he would have to wait about two years or so.

“And so I guess I would say that my use of it is through online sites, in particular, I would say JSTOR, would be – in particular is the one I think I use the most by far. A lot of the articles that I find myself wanting to research have come out four or five years previous to the time I’m looking for them. So if I recall how JSTOR works, they don’t put anything up on there usually within a

three, or four, or five-year time span, so after that it's on there, and so it's so easy to obtain articles that way." (Ph32)

While the other PhilSci participants valued mostly immediate and current research in relation to open access, scholar Ph32 has an expanded view of open access that includes delayed open access (of 3-5 years) through JSTOR. For Ph32 JSTOR augments his scholarly research process and complements PhilSci as a perceived open access repository (due to JSTOR's reduced rate of acquisition by institutions).

"Other areas, I know I have friends who have websites that they put their papers up on, that's another source I've used. And I'm finding more and more – this is something I'm still sorta getting know – is that that are archives where sort of ancient documents are put up. And those are – sometimes, I think you have to get password through the school to use them, and I haven't done as much of that, and it's something I'm gonna be doing I think in the next several years. That's also becoming a big area [having an easy access to older manuscripts by scanning them and putting them online]." (Ph32)

This is another example of scholars' Home Pages ("friends who have websites") being perceived as valuable resource and used as access tools where a scholar would go to find the latest research by a specific author. In addition to Home Pages performing on Ph32's search process, more important for Ph32 is the ability to access scanned copies of older manuscripts, where open access repositories act as translating actors that perform across time (making available very old manuscripts in electronic format) and across space (making them widely available).

"I would say, well, J – that's a good question. I see the PhilSci Archive more as papers and ideas in the making because a lot of the papers that are put up – it depends on the person. Myself, I put papers there, and a lot of times, they're rough drafts. They're – I know the ones that I'm gonna go back and re-change, and I just put them up there because I want to certain people to know what I'm working on. Also in some sense it – I feel like I put it up there and then it's kinda claiming these ideas as my own. Other people I think put it up – put only papers up there that they're certain are going to be published more or less in that same format, although I'm not quite sure of that, you know, how often that is." (Ph32)

Ph32 perceives PhilSci as a place where scholars exchange their early thoughts (and claim credit), much like they would in a conference. Therefore, PhilSci performs on scholars' research process by realigning it to ensure PhilSci is visited during the initial stages of knowledge production. It also performs on scholar's individual knowledge network by making new and sometimes unpublished knowledge visible, discoverable and accessible.

“So I see it more as almost like a worksite in many ways for a lot of people because a lot of people – a lot of those articles sometimes will never be published that people put up on the PhilSci Archive. Some are, or definitely will be coming out, so that's what I like about PhilSci Archive. It's kind of like a – I don't know what would be the word. It's sort of like state of the art for all the people who are into it, like what they're working on, whereas JSTOR of course is things that have been published and usually published three or four years past, so that's really almost just like a repository for past articles. ” (Ph32)

“You know, it's really convenient in that regard, but PhilSci Archive I really see more as kind of like – it's kinda like a place where we kinda swap our new ideas or we try them out. And so it has a little bit more of a development – developing aspect to it. ” (Ph32)

PhilSci is also perceived by Ph32 as collaboratory and a place where scholars of philosophy of science share their current work in progress and swap their current ideas informally (similarly to mingling of scholars at conferences and exchanging ideas outside of official venues) so they can develop their ideas further. Therefore, PhilSci properties of openness and integration of latest research findings or works in progress introduce scholars to ideas that otherwise would be delayed by about two years. In addition, it introduces scholars to pre-prints and the ideas therein that may never make it to commercial journals.

“I mean I think there was a time when I used to spend more time there, and then you would actually have to go look up the journal and pull it off the shelf, and then go photocopy it, which is really time consuming and a pain in the butt. And

sometimes, it would be – there would be a missing copy somewhere. And so it has really helped to make it much more efficient, and also it has encouraged that kind of browsing mentality, where I'll – you know, say I've just gotten an article on JSTOR and I'm like, "Hm, let me see what's happened the year after that," and I'll go searching through. And then since you're not photocopying and spending the money, it's so much easier to say, "Well, I'll get a copy of that as well." " (Ph32)

The availability of open access has performed on Ph32's personal research process by enabling him to use more of his time to do research instead of visiting the physical library and performing tedious time consuming searches. The availability of many articles in the same place is perceived by Ph32 to enable what he calls a "browsing mentality" by complementing JSTOR's perceived delayed open access with PhilSci's more immediate and forthcoming research. The availability of articles in electronic format vs. having to photocopy them has also performed on Ph32 knowledge production process by further enabling him to do more research by reducing the cost of acquiring new articles.

"So I think it has really helped the industry a great deal in terms of the ease of doing research and getting material. And now, like I said, that even though I really haven't quite used it as much as I plan to, the fact that all this historical stuff in terms of original manuscripts being photocopied and put up on the website. And the ability, then, to do searches through them for particular words or whatever, is going to make that kind of a really nitty gritty historical research so much easier." (Ph32)

"Because like I said – I mean I mentioned before, there was a time when you actually had to go to that library and stay there, which is an incredibly expensive option, and for many people they just can't do it. And so as these things become more available online, it's really gonna make doing historical research – which is really quite a different thing than say writing an article on the state of a philosophy problem over the last 30 years, which is really pretty much you just look at articles that other people have written. But it's even gonna make this more historical work be so much easier. I think that's gonna be a huge change, you know." (Ph32)

PhilSci is also perceived as making scholar's research process more effective and efficient, especially in searching, discovering and accessing relevant materials. Another value perceived by Ph32 is that as many historical documents are being scanned and put in PhilSci, it will enable scholars to conduct historical research. Therefore, PhilSci is seen as enabler and instigator of doing historical research that otherwise would be hard to accomplish for this specific scholar. PhilSci has the ability to perform on scholar's knowledge network by introducing him to artifacts enabling him to attempt new research directions.

"Yeah, that's more interesting. I would actually say it's – it really has helped in a couple cases because a couple of times, I've been able to through online access to get photocopy versions of say like of an old text, and it was at a point where maybe in an article where I wasn't quite sure how to advance. And if I really would have had to do an incredible research for it, maybe I just wouldn't have done it. " (Ph32)

Ph32 perceives that PhilSci enables research by speeding up the small process steps that are enacted as part of scholar's knowledge production process, and in this specific case extends the archive of working documents to increase historical depth.

"I certainly think it's helped the quality [of the research process] for that reason because when things are that easy to check, or to do, or do look up an extra issue, it – I think change is the way you approach certain things then, or over the overall project. You might – whereas before if it was a real hassle to locate something, you might just have said, "Okay, well in my project I'm just gonna have to narrow my bounds of it," you know. " (Ph32)

For Ph32, the ability to find articles needed for his research quickly and easily is very important, perceived to have increased the quality of the research process. PhilSci thus encourages Ph32 to continue research on ideas and projects that otherwise might have been hard to conduct if the materials needed for their further investigation and continuation were hard to obtain.

“Yeah, that I certainly – that’s been one of the nice things I think in that it’s kept people more connected on what people are doing. So yeah, I would say that’s been a great – I mean that certainly changed and made things – what would be the word for it? It’s a new way to think about the profession. It gives you sort of a new – it’s sort of a nice tool to know where information is being sort of exchanged, whereas before like I said, it would have to be word of mouth, maybe at a conference. It’s given sort of a site that you can go to to look for things.” (Ph32)

For Ph32, PhilSci is also a place, a search destination for scholars’ where they can find and participate in the latest discourse in philosophy of science. PhilSci emerges with the performative ability to complement conferences and informal communication between scholars (“word of mouth”), as a result of which scholar Ph32 feels more connected to other scholars.

*Impact on scholarly output (Ph32)*

As it has been mentioned earlier, Ph32 perceives that the early and quick availability of current research in the philosophy of science enables scholars to be up to date with latest research findings that can be used in their research as well. Further, PhilSci is perceived as structure with similar capabilities as conferences or even as digital extensions of in person conferences that enable scholars to share their ideas and thoughts in various stages of development and also learn from each other, performing on scholars’ local and immediate knowledge network, impacting their knowledge output.

“Yeah. Yeah, that is an interesting question. I’m not quite sure how to answer – I probably – it would have been a lot longer to do a lot of things I think. And maybe I wouldn’t have been able to find things which were really useful for the – and result of the work I was doing. And so that’s a tough kind of hypothetical one, but I would say that – I would say that it probably would’ve slowed it down a great deal and potentially weakened the end result because you just wouldn’t have as – you wouldn’t have been able to keep up on what other people were doing maybe quite as easily.” (Ph32)

“And so I would have had to order inter-library loan and gotten paper copies, and that would have definitely slowed me up, so I would say that it’s definitely improved the quality of the work I do and the ease of doing it. ” (Ph32)

P32 perceived that the time a scholar saves by not going to physical libraries and having access to the latest research findings enables him to concentrate more on the actual research, and therefore the quality of the produced article is better.

“And so I think there’s a sense in which it’s really opened up the ability to do more. And so – yeah, so and – and especially given that in the future I wanna do more looking at original documents, I think then it’s really gonna be crucial because I’m not gonna have the ability to go to some school and look at manuscripts. I may still have to in certain cases, but since more and more stuff is becoming available. But I’ll – you know, but one of the other things sometimes is that you might see something online through being able to have access to these manuscripts that will just change the way you think about a problem. And if you hadn’t seen it, you don’t know what would have happened. ” (Ph32)

Ph32 perceived that there is a relationship between ease of access and having access to the latest research and the quality of the knowledge artifact produced. Thus, these properties of open access, ease of access and early access of research findings, have performative ability in the scope a scholar puts around a specific research projects, impacting the scholarly output.

“Yeah, it’s very serendipitous. Yeah, you don’t – I mean there’s always the possibility that you may just look at something and say, “Wow, you know that’s –” and that really changes your thinking about a problem. And if it wasn’t for the online access, you might not have had that moment, so I mean it’s clear – it clearly I think is – you know, made it a much better situation. ” (Ph32)

As with some of the other scholars, Ph32 also perceives PhilSci as enabler of serendipitous discovery by bringing together different sources into an integrated information environment.

*Integration with scholarly context (Ph32)*

“But I think that’s changing because it has some big advocates in the philosophy of science community that think that this is the way we should go, and that as long as there’s referees and the quality is really high in terms of the articles that are put on it, then what’s the difference? And that’s a very good argument.” (Ph32)

“And I think actually put like a PhilSci link there, and I’ve seen PhilSci links in other journals – on published articles in journals where the link – so yeah, so I mean there is a sense in which even though I always see it less as a place – I don’t really see it quite as a publishing place. I see it as more of like an information kind of –” (Ph32)

Although Ph32 perceives that PhilSci is valuable and it has a positive impact on scholarship in the philosophy of science, it is not appropriated and used equally across by the scholars. Although individual scholars may have integrated PhilSci in their knowledge production process, other scholars do not see it as such yet. Scholar Ph32 perceived the links to PhilSci in articles as proxies to the knowledge articles published elsewhere (i.e., in scholarly journals). In a sense, the links to PhilSci in scholarly journals are meta-data (references) that point to a collective proxy (PhilSci) that points to journal articles.

*Democratization of the scholarly discourse (Ph32)*

“Yeah. I would say it’s clearly helped me a great deal because at the small school I’m at, they just don’t have many – they don’t have many journals that they keep a paper copy there. And so I would have had to order inter-library loan and gotten paper copies, and that would have definitely slowed me up, so I would say that it’s definitely improved the quality of the work I do and the ease of doing it. And so I definitely would say I’ve benefited from it. In terms of how easy they are to use online, I think the first time you – couple times you always use it, it’s kind of slow because you don’t quite know how to navigate through it.” (Ph32)



Ph32 perceives that without PhilSci he may not have been also able to participate successfully in the scholarly discourse in the discipline. He teaches at a small school that does not have access to many commercial journals required for research. Thus, PhilSci has enabled Ph32 to participate in the scholarly exchange by democratizing the scholarly discourse.

*Participant Ph32 summary in matrix form*

As it has been already described, Table 17 is a brief summary contextualizing the properties of PhilSci and the related tool and resources as experienced and perceived by Ph32, in relation to the four themes.

Table 17: Summary of OA properties and their relation to the four themes as perceived by Ph32

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>Ph32</b>	<p><b>Openness, time (early access)</b> → enables access to <b>research as it happens</b>, instead of waiting for two years</p> <p><b>Open access (via Home Pages)</b> → find the <b>latest research</b> (research process realigned)</p> <p><b>Openness, integration</b> → PhilSci with the role of a <b>conference, a collaborative tool, exchange</b> ideas and thoughts,</p> <p><b>Openness, integration</b> → <b>complementing the conference structure</b></p> <p><b>Time (faster, quicker access)</b> → able to <b>do more</b></p>	<p><b>Time (faster, quicker access)</b> → <b>up-to-date</b> with <b>latest</b> research</p> <p><b>Time (faster, quicker access)</b> → able to <b>do more</b> research rather than spend time in the physical library, <b>better articles</b> produced</p> <p><b>Openness, integration</b> → PhilSci with the role of a <b>conference, collaboratory to exchange</b> ideas and thoughts, new and <b>unpublished knowledge may enter</b> the scholarly ecosystem</p> <p><b>Easy access, time (latest research)</b> → <b>better quality</b> of articles</p>	<p><b>Home Pages</b> of authors as <b>access tools</b> for open access</p> <p><b>Google</b> used as access tool</p> <p><b>PhilSci integrated in the scholarly context of this particular scholar</b>, not so across the discipline</p>	<p><b>Openness, integration, inclusiveness</b> → enables scholars from <b>smaller institutions to participate</b> into the scholarly discover, that otherwise would not have been able to</p>

	<p><b>research</b> rather than spend time in the physical library</p> <p><b>Openness, integration (of different sources)</b> → enables <b>browsing mentality</b>, article seen that otherwise would not have been, <b>serendipitous</b></p> <p><b>Open access</b> → life easier for researcher in <b>finding, discovering, accessing articles</b></p> <p><b>Openness, time</b> (faster), <b>inclusion</b> (scanned older document) → <b>new type of research enabled/encouraged, knowledge enters the ecosystem</b> that otherwise would not have entered</p> <p><b>Open access</b> → <b>speeding up research, enabling</b> research to be done that otherwise would not have been time for</p>	<p><b>Open access</b> → enables <b>better framing of research problems</b> (knowledge network realigned)</p> <p><b>Openness, integration</b> → <b>serendipitous discoveries, better quality</b></p> <p><b>Openness, time</b> (faster), <b>inclusion</b> (scanned older document) → <b>new type of research enabled/encouraged, knowledge enters the ecosystem</b> that otherwise would not have entered</p> <p><b>Open access</b> → <b>speeding up research, enabling</b> research to be done that otherwise would not have been time for (research process and knowledge networks realigned)</p>		
--	---	---	--	--

## **Chapter 10. Comparing and contrasting the lived experiences and perceptions amongst and between the arXiv scientists and PhilSci scholars**

In this chapter, the researchers within the two groups are compared and contrasted in order to understand the range of their individual lived experiences and perceptions. Further, the lived experiences and perceptions of the arXiv scientists and the PhilSci scholars are interpreted in terms of disciplinary cultures and contrasted with each other along their properties (shown in Table 18) and dimensions that define the four themes. The performative capabilities of OA are interpreted through the identified OA properties as they relate to the four themes.

### *10.1 arXiv scientists – group level experience*

As a result of the analysis and interpretation in Chapter 8 that focuses on each arXiv scientist (as summarized in the corresponding tables for each participant), Table 18 shows an aggregated summary of the perceived open access properties of arXiv and ADS for the group of arXiv scientists, as well as the perceived properties related to open access in general, and their relation to the four themes.

An expanded and a more detailed version of Table 18 is shown in Appendix I (see Table I1). The table shows the relationship of the main perceived properties related to open access as related to the four themes, the properties of the impacted actors such as elements of researchers' knowledge production contexts, and also the properties of the repositories across the technological and organizational layers. Table I1 is used in this section and together with Table I2 (for PhilSci scholars), it is used in subsequent

sections when comparing and contrasting the lived experiences and perceptions between arXiv scientists and PhilSci scholars.

Table 18: Synthesized list of properties for each theme as perceived by the arXiv scientists

	<b>Impact on scholarly process</b>	<b>Impact on scholarly output</b>	<b>Integration with scholarly context</b>	<b>Democratization of the scholarly discourse</b>
<b>arXiv scientists</b>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/early/quick access</li> <li>- space/wider distribution</li> <li>- integration</li> <li>- linking</li> <li>- inclusiveness</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/early/quick access</li> <li>- space/wider distribution</li> <li>- integration</li> <li>- linking</li> <li>- inclusiveness</li> <li>- access to latest research</li> <li>- sub-categorization a challenge</li> <li>- trust</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- inclusiveness</li> <li>- early access to latest research</li> <li>- part of everyday life</li> <li>- trust</li> <li>- central role</li> <li>- completeness</li> <li>- total access</li> <li>- related to pre-print culture</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- inclusiveness</li> <li>- unpublished knowledge</li> <li>- smaller institutions can participate</li> <li>- bypass peer-review</li> <li>- enable new entrants</li> </ul>

For the arXiv scientists, arXiv and ADS have emerged with a central role in their discipline and their knowledge production process. These open access resources and access tools are integral in their everyday scholarly activities because they are trusted as proxies to the disciplinary knowledge network that provide an almost complete view of the scholarly landscape and the latest research findings. This is achieved by arXiv and ADS being inclusive with respect to the completeness of materials included. For these researchers there is no excuse not to be aware of the latest research findings because they can easily and quickly be found in arXiv or via ADS.

The dynamics of the performative abilities of the open access repository and the access tool are described through the implication of their properties (as they emerged from the analysis and interpretation based on scientists' lived experiences and perception) on the four themes. Here, the nature of the dynamics between scientists'

knowledge production process, scientists' knowledge network and the broader context, and the open access properties, are addressed. The nature, emergence and construction of the open access properties related to arXiv and ADS across the organizational and technological layers, and their contextualization and interrelatedness with researchers' lived experiences will be addressed in section 11.3 after the elements and properties of the organizational and the technological layers are analyzed and interpreted.

Before the implications of the perceived properties on each of the themes are presented, the properties themselves are first described, emphasizing the difference of how some of the more important properties are perceived differently by individual scientists.

The main properties that have been perceived to have an impact on scientists' production process are openness, time, space, integration, linking, and inclusiveness. The most broad, comprehensive and encompassing is the property of "openness". It is perceived to have implications for all four themes, and it is the primary property of the open access tools and the OA repositories. It is also perceived as an enabler for integration, linking and inclusiveness.

However, what do arXiv scientists mean by openness? Two distinct categories of strong perceptions amongst the arXiv scientists emerge about the role of openness: a) remover of barriers to entry, and b) enabler of new interactions. As a remover of barriers to entry, openness has enabled the participation in scholarly communication of scientists that otherwise might not have been able to participate due to cost of access to commercial journals; and openness has also enabled the entrance of unpublished knowledge artifacts (usually in the form of pre-prints) into the disciplinary knowledge

network for use freely by researchers from around the world. As an enabler of new interactions, openness has been perceived as enabler of interactions at the technical level (of linking and integration between resources and systems), and social level (enabler of collaboration between astronomers as a collective). The arXiv scientists articulated clearly that openness increases articles' visibility, discoverability, and accessibility. Thus, openness emerges as a property that has performed on the disciplinary knowledge network by making the knowledge artifacts more visible, discoverable and accessible for local use by individual scientists.

With respect to the property of time, the arXiv scientists perceive that arXiv and ADS have enabled quicker access to the latest research, faster distribution of pre-prints, and much earlier entrance of knowledge into the disciplinary knowledge network that would otherwise be possible (achieved by using knowledge found in pre-prints even before they are published in their respective journals). Further, the property of space has been perceived as enabling broader access and distribution of knowledge artifacts. Being location independent was another property that emerged related to open accessibility—scientists can access arXiv from anywhere.

Beyond the main properties of increased openness, shortened time to access, and broader access to knowledge artifacts, the arXiv scientists turn inward by articulating the internal organizational and technological properties of arXiv and ADS, such as inclusiveness of knowledge materials from many different sources, integration between arXiv and ADS, and the linking at technical level. arXiv scientists further perceive that due to arXiv's and ADS's openness and inclusiveness of broad range of materials, arXiv and ADS have over time become integrated with the disciplinary knowledge production

context. Researchers are expected to use them in their research process and contribute therein. With respect to inclusiveness, arXiv scientists perceive that arXiv and ADS are almost complete proxies to the disciplinary knowledge network in astronomy and astrophysics, including here access and visibility to raw astronomical data, with very few disciplinary journals missing. In addition, the inclusiveness of old scanned journals was perceived to be of great value for knowledge production for some of the scientists. The linking between arXiv and ADS at technology level (exchange of meta-data, ADS's ability to fully search arXiv, automatic citation linking, etc) in conjunction with the open access to pre-prints and post-prints, emerges as one of the performative abilities that has increased article visibility, discoverability and accessibility.

Next, the roles and values of these properties as they are implicated in the four themes are discussed.

#### *Impact on scholarly process*

As it has been already discussed, some of the perceived properties of arXiv and ADS that have enabled such integration with scientists' knowledge production process are: openness, interoperability and linking with other systems, the perception of an almost complete access to all scholarly materials available in astronomy and astrophysics, fast and quick access to the scholarly material, and integration with other open access tools.

These properties have performed on scientists' knowledge production process by realigning scientists' research process, article search steps, as well as the writing stages. Some of the realignment of process is related to being able to find the latest research findings faster through a normalized approach (via ADS or directly into arXiv, instead of many individual journal or library portals), enabling the scientists to spend more time

on doing research (pondering over the nature and scope of the problems) instead of spending a lot of time searching for articles. The inclusiveness of many different sources into an aggregated source such as arXiv and accessible via its own portal or via ADS, has realigned the search process by enabling the emergence of new structures in the form of new tools that otherwise would not have been feasible. For example, due to arXiv's open interface and open metadata exchange standards, ADS' search and browse capabilities are linked with arXiv at bibliography and citation level. As a result, arXiv scientists can traverse with ease from one article to the articles in its bibliography through clicks, and they can easily find who has cited a specific article and thus assess citation impact. As a result, ADS has become one of the main tools for citation searches, and it is also used by tenure committees during the tenure assessment process.

Thus, not only have arXiv and ADS realigned the knowledge production processes of individual scientists by making them more efficient, they have performed on the disciplinary norms by enabling the emergence of a new pattern of use at disciplinary level such as the use of ADS by tenure committees. At individual level the scientist use ADS in their knowledge work (citation searches), while at departmental or institutional level the scientists use ADS for citation impact assessment, both enabled by ADS' bibliographic linking capabilities. From ANT's perspective, the scientists themselves are the conduits through which ADS has been appropriated for use in the tenure assessment process. In this case, the appropriation process comprises of translating individual and localized capabilities for communal good at group level, by identifying the values and benefits that emerge from aggregating individual level values.



Further, because of the new interactions at social level, and the entrance of scientists from smaller institutions into the scholarly communication process, and the location independent property of arXiv and ADS, scientists have been able to collaborate more widely and quickly. Open access properties have thus translated scientists' knowledge production process by "shrinking" time (more research can be done in less time) and by "shrinking" space (more knowledge artifacts can be consulted from a broader and distributed range of sources), providing an immediate knowledge network for local use. Whereas before open access scientists had to spend more time across a number of internal portals within their library as well as journal portals, unsure whether they have identified all relevant articles for the research problem at hand, open access has enabled the emergence of an efficient and effective local and immediate individual knowledge production context where the individual knowledge network is a subset from the disciplinary knowledge network. Using the language of ANT, arXiv and ADS can be perceived as a funnel that translates the global to local. On one end, they aggregate a vast and dispersed disciplinary knowledge network, and on the other end, they enable effective and efficient processes that can be enacted by each individual scientist to produce a local knowledge network based on the criteria for specific research problem.

#### *Impact on scholarly output*

In addition to the impact on their knowledge production process, arXiv scientists have perceived impact on their scholarly output. The properties that have been perceived to have an impact on scientists' knowledge output are: openness, immediate access to latest research, access to broad base of resources, an almost seamless integration and

link between arXiv and ADS and raw data repositories, inclusiveness of different sources, sub-categorization of the corpus of materials available in arXiv, and trust. By performing on scientists' production processes, enabling articles to enter the disciplinary knowledge network earlier than otherwise would have been possible, the scientists can produce articles that include knowledge that would not have been included at that particular time. Further, ADS and arXiv perform on scientists' knowledge output by enabling them to use knowledge from pre-prints that otherwise might not have entered their scanable knowledge network. The availability of many articles (pre-prints and post-prints) and conference publication in one location, searchable and browseable via the same venue, have also enabled serendipitous discovery, triggering new ideas and research problems, thus enabling scientists to work with a broader archive of knowledge artifacts.

#### *Integration with scholarly context*

The deep integration of arXiv and ADS in everyday scholarly life of researchers is perceived to have performed on scientists' knowledge production process, by enabling them to find and discover a wide range of materials needed for knowledge production much easier and quicker, without any serious impediment to accessing the full text, as a result of which scientists rarely visit the commercial journals. The integration with everyday life and the perceived completeness, inclusiveness and trust in arXiv and ADS through which the scientists can access the latest research findings immediately after they are conducted, has also realigned the knowledge networks of each of them by enabling them to produce better quality articles, work on research problems that would

have otherwise been unrecognized, and have acted as triggers for research problems and ideas due to the proximity of multitude of scholarly resources in one place. Some of the scientists mentioned serendipitous discovery as an important outcome of their interaction with arXiv and ADS. While the arXiv scientists perceive arXiv and ADS as central and integrated with their daily scholarly activities, they still perceive that the commercial journals are important due to their function as quality and relevance judges via the peer-review process. Thus, the increased visibility, discoverability and accessibility of articles via central open access resources such as arXiv and ADS have performed on scientists' knowledge production context by aggregating the disciplinary knowledge network and making it visible and available for local use by the scientists. Both arXiv and ADS play the role of the aggregators by translating a distributed ecosystem of knowledge artifacts into a centrally accessible one, where specific search capabilities (such as browsing, searching, bibliographic linking) act as filters that translate multitude of search interfaces into a uniform search interface.

Further, the traditional paper based pre-print culture in the physics scholarly community is perceived by the arXiv scientists as having impacted the appropriation of electronic pre-prints and electronic access tools with a central and critical role in the digital realm, manifested as arXiv and ADS. The accepted practice of paper based pre-print sharing has performed as an enabler of open access, inscribing its properties of openness and sharing onto the electronic scholarly publishing practices by replacing the paper based distribution process with digital distribution.

### *Democratization of the scholarly discourse*

The arXiv scientists largely also perceived that the availability of arXiv and ADS as open access resources that enable more efficient and effective research process and more complete access to the disciplinary knowledge network, has democratized the scholarly landscape by enabling scientists from smaller institutions to enter the scholarly discourse in astronomy and astrophysics. Few of the scientists perceived that the availability of pre-prints via arXiv and ADS and the integration of these open access resources with everyday scholarly life have also democratized the disciplinary knowledge network by enabling articles to bypass the peer-review process and enter the disciplinary knowledge ecosystem. These few scientists perceived that often good enough articles do not make it through the peer-review process and do not get published in commercial journals because of administrative and logistic restrictions such as journal page numbers, peer-reviewers' stance on certain research problems, the fuzzy logic with which reviewers decide what to accept and what not. Although this perception was not expressed by all participants, it is interesting to note that arXiv and ADS may be able to perform on the disciplinary knowledge network by enabling unpublished knowledge to enter the knowledge ecosystem.

### *10.2 Differentiating and unique experiences amongst the arXiv scientists*

The following are some unique and differentiating lived experiences and perceptions amongst the arXiv scientists.

Although one of the participants in the study (scientist A18) uses arXiv to access the astronomy related knowledge artifacts like the rest of the scientists, she uses CiteSeer instead of ADS that has been used by the other scientists. The use of CiteSeer by A18 is

driven by her role as an interdisciplinary researcher whose role is to build and manage software tools related to astronomy. CiteSeer is her window into “information technology and information management articles” that are not included into ADS or arXiv but she needs them for her research work that includes software development. CiteSeer performs on A18 by acting as an aggregator, translating a dispersed and distributed set of resources into a uniform access for local use, similarly to ADS, by complementing A18’s knowledge network with interdisciplinary materials.

Participant A17 perceived that the combination of arXiv and ADS enabled him to observe two articles in close proximity to each other, and thus enabled him to see a potential link between the articles that triggered a new research question. A17 further discussed this serendipitous trigger of a research problem with a colleague and they have collaborated on a few papers as a result. The interesting outcome is that A17 considers these papers tangential to his area of research and would not have otherwise engaged to address the specific problem if it was not for the serendipitous discovery instigated by arXiv. Thus, in this case the openness and availability of articles from two different journals in one place not only had impacted A17 by realigning his production process to enable collaboration with a colleague, it has also performed on A17’s knowledge network (perceived as a more complete view of the disciplinary knowledge ecosystem) by enabling him to engage in a research problem that might have stayed outside of his research interest.

For participant A19, arXiv has had even more profound impact. A19 had submitted a pre-print of her latest research in arXiv. As a result, even before she could learn whether it had been accepted for publication, based on her submission to arXiv she was invited to

speak at an important international conference. A19 perceived this as a real “aha” and “wow” moment and contributed her career advancement to arXiv. Thus, the availability of arXiv as a collaboratory for sharing locally produced knowledge at disciplinary level, even for knowledge presented in pre-prints that may not even make it through the peer-review process, has performed on A19’s research career trajectory. In this specific case, arXiv was able to reconfigure the knowledge network available to the participants in the conference, and it may further enter the individual knowledge networks of many scientists not present at the conference who may consider the knowledge artifact in their knowledge production. As a result, her research findings, once deposited in arXiv enter the global knowledge network (i.e., disciplinary knowledge ecosystem) which can be further pulled locally by scientists, independent of their location. There are two senses of “location” that have emerged here. First, in terms of power rearrangement arXiv, has positioned A19 (through her pre-print) to affect the participants at the conference by reconfiguring the knowledge network presented at the conference, more so as an invited speaker. Second, arXiv acts as a conduit through which locally produced knowledge can be distributed to a wider audience that is further aggregated with other distributed resources for further local use by other scientists where arXiv and ADS act as the “filtering” and “localizing” lens.

Scientist A20 brought forth an interesting observation from her experience over the years about a different (spatial) sense of location. She perceives that the average number of authors per paper has increased as a result of the open access resources in conjunction with the availability of electronic communication. The replacement of postal mail, phone and faxes with electronic collaboration contexts (such as e-mail, electronic workspaces,

open access repositories) has instigated a more open and collaborative culture on international level. In essence, open access, electronic communication and the availability of scholarly work in digital format have removed the barrier to collaboration by “shrinking” time and space, resulting in new social interaction amongst scientists. This in turn has enabled more researchers to work together on the same set of problems and thus share authorship on published works. This implied implication of open access to translate the nature of authorship to impacting the number of authors per paper, differs from A19’s experience relaying the story of the “morning coffee” as a new type of authorship emerging in arXiv. In both cases open access has had implication on translating the nature of the concept of authorship as a collective of researchers. In the case of the increased number of authors per paper, the authorships concept is modified with individual authors remaining visible in the final scholarly product, while with “morning coffee” a new author construct emerges with individual researchers becoming invisible in the final scholarly product.

### *10.3 PhilSci scholars – group level experience*

As with the corresponding arXiv sub-section, in this sub-section the aggregated lived experiences and perceptions amongst the PhilSci scholars are described.

As a result of the analysis and interpretation in Chapter 9 that presented the lived experiences and perception of each PhilSci scholar (as summarized in the corresponding tables for each scholar), Table 19 shows an aggregated summary of the perceived properties for the group of PhilSci scholars, related to PhilSci, as well as perceived properties related to open access in general, and their relation to the four themes.

Table 19: Synthesized list of properties for each theme as perceived by the PhilSci scholars

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>PhilSci scholars</b>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/fast/quick</li> <li>- space/wide</li> <li>- some integration with conferences</li> <li>- access to latest research</li> <li>- collaboration / collaboratory</li> <li>- Home Pages as access tools</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/quick/fast</li> <li>- space/wide</li> <li>- integration</li> <li>- early access to latest research</li> <li>- some role for invisible colleges</li> <li>- trigger of new ideas</li> <li>- some serendipity</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/quick/fast</li> <li>- space/wide</li> <li>- pre-print culture</li> <li>- collaboratory</li> <li>- limited role</li> <li>- no central role</li> <li>- Home pages and other access tools</li> </ul>	<ul style="list-style-type: none"> <li>- smaller institutions can participate</li> </ul>

A more detailed version of Table 19 is shown in Appendix I (see Table I2). It shows the relationship of the main perceived properties related to open access as related to the four themes, the properties of the impacted actors such as elements of scholars' knowledge production contexts, and also the properties of the repositories across the technological and organizational layers.

As it is shown in Table 19, similarly to arXiv, the properties of openness, time, and space are perceived as having an implication for the four themes. PhilSci scholars perceive the property of openness as remover of barriers to entry as well as an enabler of new interactions. As a remover of barriers to entry PhilSci is perceived to have enabled the entrance of scholars from institutions that otherwise might not have participated in scholarly communication due to availability of a limited knowledge network, and that the latest research can enter the knowledge networks of individual scholars much earlier rather than entering the knowledge network via the commercial journals after many months, sometimes up to two years. PhilSci makes possible for yet unpublished pre-prints (some of which might not be accepted for publication) to enter the disciplinary



knowledge network. As enabler of new collaborations, for PhilSci scholars the interactions are enabled at a social level with PhilSci perceived to be enabling new collaborations between scholars (by exposing scholars to each other's works, sometimes it is work in progress) and that it also acts as a preparatory place for conferences.

Beyond the perception of the main properties (openness, time, space) as strongly related to the availability of the PhilSci archive as an open access repository that enables scholars to find, discover and access the latest research finding faster and with ease, PhilSci scholars looked outwards by describing scholars' Home Pages (as a distributed form of open access), invisible colleges (scholars know who is working on specific problems due to their existing relationships with other scholars), and JSTOR (perceived as a delayed open access due to its low cost of acquisition for institutions; however it does not contain the most current issues published in the past 3-5 years) as complementary to PhilSci that increase the visibility, discoverability and accessibility of the articles needed for their knowledge production.

There are two broad patterns of perception for the PhilSci scholars: a) the PhilSci archive enables them to find, discover and access the latest research findings faster and with ease, and b) PhilSci has a limited role in their discipline and that they can still write their articles without PhilSci.

### *Impact on scholarly process*

With respect to the open access properties that are implicated in scholars' knowledge production process, PhilSci is perceived as enabling efficient search and research process by reducing the time it takes scholars to search, discover and access the

materials for their knowledge production work, while at the same time accessing a broader set of resources via the PhilSci archive. It also enables PhilSci scholars to immediately access the latest research findings. Further, PhilSci is strongly perceived as a collaboratory for scholars to connect and exchange ideas, and it is also perceived as a preparation place, as well as a publication place for conference papers.

### *Impact on scholarly output*

With respect to the impact on the scholarly output, PhilSci is perceived as a tool that has enabled scholars to: a) access the latest research findings right immediately after they are conducted; b) reduce their search time and thus conduct more research, c) increase serendipitous discovery (by bringing article from differences sources close proximity to each other in space and time), d) position the elements of their research problems within the landscape of existing knowledge and currently worked problems and thus avoid “reinventing the wheel” (Ph26), e) more efficiently share conference proceeding before the conference takes place, and f) share pre-prints amongst the participants in the discipline even if they are not published. The perception that PhilSci enables access to the latest research stands out more prominently than other perceived values enabled by PhilSci. This reveals that access to the latest research is perceived as an important actor in the production of new knowledge amongst the philosophers of science.

Therefore, PhilSci appears to be also perceived as important not only in enabling more efficient research process, but with potential to impact the knowledge output as a trigger of new ideas (by having access to the latest research) and potentially as an

enabler for unpublished knowledge to enter individual knowledge networks, as the participants noted in their interviews.

*Integration with scholarly context*

The main open access properties that to some degree have impacted PhilSci's integration with scholarly context are: a) the traditional paper based pre-print culture of sharing, b) as enabler of collaboration, c) fast and immediate access to latest research, d) the wider dissemination of pre-prints and post-prints, and e) the complementary role with respect to the use of scholars' personal Home Pages and invisible colleges in the knowledge production context. It would seem that the PhilSci scholars perceived PhilSci as a tool that has for the most part enabled them to engage in the same scholarly endeavor that they did even before PhilSci was available—by inscribing the paper based pre-print culture and the traditional mode of scholarship into the electronic context, with the difference that now they can do it more efficiently by enabling quick and faster access to the latest research findings. Apart from the enabled efficiency in searching, PhilSci is perceived to have a very limited role in the everyday scholarly life of philosophers of science and their discipline. It is mostly perceived as a collaboratory for exchange of ideas, a preparatory place for conferences and conference papers, and the main destination for finding the latest and up-to-date research findings in philosophy of science.

However, PhilSci scholars do not exclusively rely on PhilSci to find the latest research findings. Their research process is augmented and complemented by accessing JSTOR (perceived as an enabler of delayed open access due to low cost for institutional

subscriptions), the personal Home Pages of scholars and invisible colleges. Partially, the seeming contradiction can be explained by the perception that PhilSci is not the only source that enables the finding and accessing of the latest research. The search process is also enabled and complemented by scholar's personal Home Pages. Thus, scholars' Home Pages act as a distributed form of open access repository, without any centralized and specialized access tool. PhilSci scholars mostly use Google Scholar or Google in their research process as an access tool.

#### *Democratization of the scholarly discourse*

The concept of PhilSci as an enabler of democratization of the scholarly discourse is not very strongly perceived by the PhilSci scholars. The value as a remover of barriers to entry is however perceived very strongly by scholar Ph32, who emphasized that PhilSci is important in his information practices because he works in a small institution that does not have access to the required journals that he needs for his research.

#### *10.4 Differentiating and unique experiences amongst the PhilSci scholars*

The following are some unique and differentiating lived experiences and perceptions by specific PhilSci scholars.

Unlike the rest of the PhilSci scholars, Ph32 strongly perceives that PhilSci has enabled him to write better quality articles, thus acknowledging that PhilSci, through its property of openness and enablement of easy and fast access to the latest research findings, has realigned his knowledge production process by making a broader set of knowledge artifacts available for local use. In addition, Ph32 is the only one from the

PhilSci scholars to perceive that PhilSci has democratized the scholarly discourse in philosophy of science by enabling researchers from smaller institutions to enter the scholarly publishing process. He identifies himself as a researcher from a small university. Ph32 is also unique amongst the PhilSci scholars in that he perceived open access as a vehicle that will enable old scanned documents to be easily available to scholars across the world, and that it will encourage more original research based on the scanned manuscripts and documents.

For PhilSci scholars there was no “wow” or “aha” (perceived as extremely or uniquely helpful) such as those experienced by arXiv scientists. It is important to emphasize instead the distinction that has been made by all PhilSci scholars, most clearly articulated by Ph24 and Ph29. Apparently, they value greatly the ability to access the latest research finding immediately and without delay, and yet, the knowledge acquired by reading the latest research findings right away is acknowledged as having only a limited impact on the scholarly output. All together, PhilSci is perceived as having no impact on the disciplinary knowledge production context and that the scholars do not necessarily depend on PhilSci when writing their articles. This might be partially explained by the acknowledgement that the access to open access resources is distributed across PhilSci, scholars’ Home Pages, and the invisible colleges, where PhilSci is only one of the ways to access the latest research findings.

As it has already been mentioned, the PhilSci scholars perceive PhilSci’s value with respect to quick and fast access to the latest and cutting edge research. For example, Ph29 emphasized strongly that he values PhilSci mostly because it enables him to distribute his work to others quickly and easily, in comparison to Ph26 who perceives

value in PhilSci to understand the current ongoing research so he can make sure he works on something unique and not on research problems that others are working.

### *10.5 Comparison of lived experiences and perceptions between astronomers and philosophers and science*

The following section compares and contrasts the lived experiences and perceptions between the arXiv astronomers and philosophers of science as related to the four themes. All of the researchers had expressed perceptions related to the four themes with the following exceptions. Amongst the astronomers, although A20 perceived that arXiv has enabled collaboration with scientists from around the world, she did not explicitly elaborate about the role of arXiv with respect to “democratization” of the scholarly discourse. Amongst the philosophers of science, Ph32 was the only who explicitly articulated that the open access resources are implicated in the “democratization” of the scholarly discourse, while the rest perceived value in PhilSci’s role to enable collaboration amongst the scholars.

To aid in the analysis, the summarized tables (see Table 18 and Table 19) that represent the list of properties for each theme as perceived by the arXiv scientists and PhilSci scholars, have been put together side by side in Table 20 for comparison. It can be observed that both the arXiv scientists and the PhilSci scholars commonly perceive the value and role of open access repositories and the related access tools to emerge from their properties as: a) enablers of openness, b) enablers of early, quick and fast access to scholarly materials, and c) enablers of wider dissemination of scholarly material. Beyond these commonalities, there are distinct differences in how the arXiv scientists perceive the value of arXiv from the way the PhilSci scholars perceive PhilSci.

Table 20: Synthesized list of properties for each theme as perceived by the arXiv scientists and PhilSci scholars

	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>arXiv scientists</b>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/early/quick access</li> <li>- space/wider distribution</li> <li>- integration</li> <li>- linking</li> <li>- inclusiveness</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/early/quick access</li> <li>- space/wider distribution</li> <li>- integration</li> <li>- linking</li> <li>- inclusiveness</li> <li>- access to latest research</li> <li>- sub-categorization a challenge</li> <li>- trust</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- inclusiveness</li> <li>- early access to latest research</li> <li>- part of everyday life</li> <li>- trust</li> <li>- central role</li> <li>- completeness</li> <li>- total access</li> <li>- related to pre-print culture</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- inclusiveness of unpublished knowledge</li> <li>- smaller institutions can participate</li> <li>- bypass peer-review</li> <li>- enable new entrants</li> </ul>
<b>PhilSci group of scholars</b>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/fast/quick</li> <li>- space/wide</li> <li>- some integration with conferences</li> <li>- access to latest research</li> <li>- collaboration / collaboratory</li> <li>- Home Pages as access tools</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/quick/fast</li> <li>- space/wide</li> <li>- integration</li> <li>- early access to latest research</li> <li>- some role for invisible colleges</li> <li>- trigger of new ideas</li> <li>- some serendipity</li> </ul>	<ul style="list-style-type: none"> <li>- openness</li> <li>- time/quick/fast</li> <li>- space/wide</li> <li>- pre-print culture</li> <li>- collaboratory</li> <li>- limited role</li> <li>- no central role</li> <li>- Home pages and other access tools</li> </ul>	<ul style="list-style-type: none"> <li>- smaller institutions can participate</li> </ul>

The property of openness was ubiquitously referred throughout the interviews with the researchers and it is often used as a generic term to denote the overall value provided by the open access resources. The properties related to time such as fast, quick and early access to scholarly materials, and space such as ability to read materials from researchers from around the world, are also common between the two groups. The open nature of the open access repositories, driven by the removal of the access barriers so that pre-prints, post-prints and their meta-data can be freely accessed, has enabled open access repositories to emerge as centralized (artifacts from different sources can be

accessed through the same interface), localized (open access repositories act as filters through which the global resources can be localized for specific research problem), and integral (artifacts from different sources are brought together) socio-technological structures. What has enabled the centralized, localized and integral nature of the open access repositories is a combination of change in the property of time and space. For example, bringing materials from many different sources together is perceived as “shrinking” space, and the fast and quick access is perceived as “shrinking” time. This in turn has had an impact on the materials deposited in the open access repositories by increasing their visibility (included in abstracting and indexing services, search engine indexes, as well as the inclusion in a repository), discoverability (their location can be determined; the article might or might not be accessible) and accessibility (their content can be read by humans or machines). The increased visibility, discoverability and accessibility of articles, has in turn performed locally on researchers’ individual knowledge production contexts by enabling the knowledge artifact from the disciplinary knowledge network to enter the locally enacted individual knowledge production contexts of individual researchers. The socio-technological aspects that increase articles’ visibility, discoverability and accessibility will be discussed in detail in Chapter 11. Apart from these common properties, arXiv scientists and PhilSci scholars perceive the value and role of their corresponding open access repositories differently, as it is described in the rest of this section.

Even the property of openness is perceived to have different implications for researchers’ individual knowledge production context. arXiv scientists’ and PhilSci scholars’ understanding of openness is perceived as remover of barriers to entry and as



enabler of new interactions. With respect to the perception as a remover of barriers to entry both groups of researchers perceive that open access repositories enable new researchers to participate in the disciplinary scholarly discourse (especially those from smaller institutions) and that open access repositories enable new research findings to enter the disciplinary knowledge network immediately (as pre-prints and manuscripts) without having to rely on the commercial journals for immediate distribution. However, while arXiv scientists perceive and articulate in great detail the enablement of interactions at technological level (automatic linking within and between the resources deposited in arXiv and ADS) and social level (as enabler of collaborations between researchers from different institutions), PhilSci scholars articulate the value of the interactions as enabler mostly at the social level (enabling collaborations amongst scholars) and very little is said about the PhilSci as a performative actor except that it is a tool that helps them find the latest research findings quicker. While arXiv scientists turn inward and perceive the value of openness in the co-construction of arXiv's capabilities to enhance the search, discovery and accessibility to the knowledge artifacts by enhancing the technological features to help the scientists, PhilSci scholars turn outward and describe the interactions only at the social level, including the complementary nature of Home Pages, invisible colleges, Google Scholar and JSTOR in their search for knowledge artifacts. For the arXiv scientists the translation (i.e., aggregation of resources in central location) capabilities are inscribed into the organizational and technological layer of the open access repositories and access tools, while for the PhilSci scholars the aggregating of resources is enacted by the scholars themselves by visiting PhilSci, scholars' Home Pages, utilizing invisible colleges, and

Google search—where the PhilSci archive acts only as a partial aggregator to enact the individual knowledge networks of PhilSci scholars. Thus, the arXiv scientists perceive the combination of arXiv and ADS as central in how they articulate their individual knowledge networks as central nodes, visiting for the most part only ADS in their search process since ADS has access to all of the records in arXiv. PhilSci scholars however, perceive the open access resources as dispersed across different contexts, with arXiv providing access to the most current and up to date research, where Home Pages of scholars act as a distributed open access repository (sometimes linking to each other with manual links) and JSTOR is perceived as a delayed open access repository due to its low cost of subscription by institutions—the most current 3-5 years of articles are not included in JSTOR.

The concept of collaboration has different meaning for the two groups of scholars. For the philosopher of science PhilSci is an enabler of collaboration for exchanging ideas and thoughts, and not as much about collaborating together on joint papers. Astronomers on the other side perceive collaboration as a way to work with other scholars on joint research problems and share authorship.

In addition to the property of openness, the property of inclusiveness of many different sources in one place emerges as a critical property that defines the nature of the open access repositories, enabled by technological capabilities of interoperability and integration and by the enacted policies by the organizing and management structures about what is to be included in the repository. As it will be described in more detail in Chapter 11, arXiv (more specifically the astro-ph section relevant to astronomy and astrophysics) and ADS have been co-constructed over time to include almost all of the

research output in the discipline. Thus, arXiv and ADS have emerged as trusted location independent central actors that are perceived as a proxy to the complete body of knowledge artifacts that define astronomy and astrophysics. This has in turn performed on the knowledge production process of arXiv scientists by gradually substituting arXiv and ADS for the library and journals' portals. It has also enabled scientists to view easily and efficiently a subset of the disciplinary knowledge network, by building a localized knowledge network (through searching and downloading of pre-prints and post-prints) that is required for the production of a specific knowledge artifact. The property of inclusiveness was not very strongly perceived by the PhilSci scholars—their view of open access is more distributed in nature as we have seen earlier.

In addition to the properties described so far in this section, a more comprehensive description of actors and their perceived properties that emerged from the interviews is shown in Table 21. Each property is then categorized for the two groups of participants along its dimension. Here, few additional properties as perceived by the two groups of participants are compared and contrasted to help in understanding the performative capabilities of the open access repositories and access tools for participants' knowledge production process and knowledge network.

Table 21: Model comparing lived experiences and perceptions

Actors	Actors' properties as perceived by the participants	Dimensions of the properties as acknowledged by arXiv scientists and PhilSci scholars	
		arXiv scientists	PhilSci scholars
Access tools and repository	Linked and integrated	High	None
	Enable immediate access to latest research	High	High (not exclusive)
	Enable more efficient research process	High	High

Actors	Actors' properties as perceived by the participants	Dimensions of the properties as acknowledged by arXiv scientists and PhilSci scholars	
		arXiv scientists	PhilSci scholars
	Trusted	High	Low
	Relevant for the tenure process	High	None
	Enable unpublished knowledge to enter the knowledge ecosystem	Medium	Low
	Enhance article visibility, discoverability and accessibility	High	High
	Researchers no longer visit commercial journals	High	Low
	Complement the peer-review process	High	Low
	Increase citation impact	High	High
Repository	Impacted by the pre-print culture	High	Medium
	Part of researchers' everyday life	High	Very low
	Provides complete access to disciplinary knowledge	High	Low (only latest research findings)
	Used as collaboratory for sharing ideas	High	High
	Important	High	Low
	Enables open access	High	High
	Enables serendipity	High	Medium
Access tools	Specialized	High	Low (generic tools)
Raw data repositories and access tool	Relevant	High	None
Scholars' Home Pages	Relevant for search process	Low	Medium
	Provide access to latest research	Low	High

Source: Interviews with the researchers.

Because of the acquired trust and comprehensive inclusiveness of materials from multitude of sources, arXiv and ADS are perceived as being very relevant for the tenure process, while PhilSci is perceived as having no relevance for the tenure process amongst the philosophers of science. Tenure committees in astronomy and astrophysics heavily rely on ADS to assess scientists' citation impacts as noted earlier. This is an example of how the openness property of the open access tools and open access repositories has gradually over time performed on arXiv and ADS, by enabling,

aggregating and translating the properties of linking, integration and inclusiveness, and inscribing them into the tenure process to be utilized by the tenure committees.

Few more differences from Table 21, such as the relevance of raw data repositories and specialized access tools, are perceived as relevant by either the arXiv or the PhilSci participants for their knowledge production context. They will be described in Chapter 12 after the organizational and technological properties of the open access repositories and the access tools are analyzed first.

## **Chapter 11. The role of the open access repositories and the access tools in researchers' knowledge production processes**

The lived experiences layer that was discussed and analyzed in Chapter 7 through Chapter 10 provides only a partial account about the nature of the interaction dynamic between the participants and the open access repositories. To understand and describe how the access tools and open access repositories realign participants' knowledge production contexts, and the degree to which they have become integral actors in the disciplinary knowledge production context, the next two sections address the implications of the organizational and technological layers, based on the documentary evidence that describes the organizational and technological elements of the OA repositories and the access tool. The co-construction dynamics between the organizational layers (section 11.1) and the technological layers (section 11.2) is analyzed in section 11.3. The implications of the technological and organization layers on articles' visibility, discoverability and accessibility are presented in section 11.4.

The commonalities and differences between the OA repositories and access tools used by the two groups of participants are interpreted at each layer (organizational and technological) and between the two layers in order to explicate the nature of the co-construction dynamics between the layers. By delineating the organizational and technological layers clearly, the analysis will explicate the socio-technological dynamics of the open access repositories and the access tools for the two groups of participants, and thus provide a link for understanding the nature of the co-construction dynamics

between the lived experiences and perceptions of astronomers and philosophers of science.

### *11.1 The organizational perspective*

In this section, the arXiv and PhilSci repositories and their respective access tools, protocols, and standards are described as purposefully organized structures of technical capabilities and functions that make them what they are: socio-technological information environments operating as archives of scholarly materials that are intended to be used by researchers in their knowledge production. The difference from the technological layer is that this section elaborates on how the technological capabilities and functions are intentionally aligned and brought together with a specific goal in mind: to produce the repositories. To clarify further with an example, a discussion of EPrints's software features and capabilities belongs in the technological layer. Discussion of PhilSci belongs in the organizational layer, where PhilSci can be described as a purposefully organized structure composed of the EPrints software, the protocols and standards, policies, and the intent and the appropriation of the features and capabilities towards a certain goal—to enable the scholarly community of philosophers of science to share amongst themselves the scholarly materials that can be used in their knowledge production.

At this level, PhilSci, arXiv and ADS are further analyzed and interpreted with respect to their ability to realign various elements that constitute researchers' individual knowledge production contexts. The evidence is extracted from the documents that describe the organizational properties and intended roles and values of PhilSci, arXiv, and ADS.

The organizational properties of the OA repositories and their corresponding access tools are analyzed and interpreted first (arXiv, ADS, PhilSci), followed by the organizational properties of the standards and protocols utilized by the repositories and the access tools.

#### 11.1.1 arXiv (organizational)

“Started in August 1991, arXiv.org (formerly xxx.lanl.gov) is a highly-automated electronic archive and distribution server for research articles. Covered areas include physics, mathematics, computer science, nonlinear sciences, quantitative biology and statistics“ (About arXiv).

“arXiv is maintained and operated by the Cornell University Library with guidance from the arXiv Advisory Board and the help of numerous subject moderators” (About arXiv)

As an organizational structure, arXiv was initiated as a single person project known as xxx.lanl.gov with the intention to enable researchers in the physics and related natural science disciplines to share their pre-prints, post-prints and other knowledge artifacts. It is managed and maintained by the Cornell University Library, and it is run by the Advisory Board.

“arXiv submissions are meant to be available in perpetuity. Thus, arXiv has high technical standards for the files that are submitted. The submission process begins with the preparation of valid metadata for the paper and continues through a verification process whereby the files uploaded are checked for certain problems. Valid upload formats for the primary text file include (La)TeX, PostScript, PDF, and HTML” (arXiv Primer).

“As an electronic archive, arXiv makes a commitment to provide persistent access to all announced submissions. arXiv is thus maintained with a focus on the perpetual availability of submissions. This is accomplished in part by controlling the types of files that can be uploaded to arXiv, as well as restricting changes that can be made after submissions are announced. In order to preserve the scholarly record, submissions are not removed from arXiv after they have been announced” (arXiv Primer).



The intended goal to offer persistent access to all submitted artifacts in perpetuity has been achieved through the normalization and control of the type of files it can support, by using and appropriating the relevant technology to support a set of standardized file types and formats. In addition to the appropriation and adoption of technological features, a policy of non-removal of already submitted materials is instituted to ensure persistency with long-term goal in mind. This ensures long-term availability and access to the knowledge artifacts that can be easily referenced and relied upon in the future.

“arXiv is proud to be able to offer such a large collection of scholarly work in a single location, without any fees and with support for users around the world” (arXiv Primer).

The intended mission to offer a large collection of scholarly materials from a single collection, for free, to interested individual around the world, is congruent with arXiv scientists’ perception that such resource can impact scientists’ knowledge production process and knowledge networks. This has the potential to democratize the disciplinary scholarly communication by realigning the scholarly knowledge production processes to enable the participation by scientists and institution that would have otherwise not been able to participate due to economic challenges. Further, due to the democratization of the production process, as was perceived and expressed by the researchers in their interviews, there is a potential for new knowledge to enter the disciplinary knowledge ecosystem. This would be the knowledge that otherwise would not have been produced because of the barriers to entry, or the knowledge that is produced through serendipitous discoveries enabled by the availability of many different sources in one location at the same time.

“arXiv supplements the traditional publication system by providing immediate dissemination and open access to scholarly articles (which often later appear in conventional journals)” (arXiv Primer).

arXiv is intended to be a complementary service for scientists alongside the availability of the commercial journals, by enabling immediate dissemination of research findings without intervening in the established publication infrastructure. This enables new research findings to enter scientists’ knowledge networks earlier than if relied on the commercial journals only.

“Users can retrieve papers from arXiv via the web interface. Registered authors may use our web interface to submit their articles to arXiv. Authors can also update their submissions if they choose, though previous versions remain available” (About arXiv).

“Listings of newly submitted articles in areas of interest are available via the web interface, via RSS feeds, and by subscription to automatic email alerts” (About arXiv).

“As a user, your primary interactions with arXiv will likely be to browse and view articles, and perhaps also to submit articles. Submitting articles requires registration as a user, whereas browsing can be done without registration, on either the main arXiv.org site or on any of arXiv's mirror sites around the world” (arXiv Primer).

The interface with the scholarly work is achieved through its portal that enables users to retrieve articles of interest via searching, browsing, or via e-mail alerts. The different means of access are intended to enable scientists to adapt the use of arXiv to their individual search and research information practices.

“arXiv is an openly accessible, moderated repository for scholarly articles in specific scientific disciplines. Material submitted to arXiv is expected to be of interest, relevance, and value to those disciplines. arXiv reserves the right to reject or reclassify any submission. Submissions are reviewed by expert moderators to verify that they are topical and refereeable scientific contributions that follow accepted standards of scholarly communication (as exemplified by conventional journal articles)” (arXiv Primer).

“It is important to note, however, that arXiv is not a repository for otherwise unpublishable material, nor is it a refereed publication venue. The moderation process is essential to ensuring that submissions are of value to the arXiv communities, but there is also a limit to the ability of administrators and moderators to provide feedback on submissions that are determined to be inappropriate for arXiv” (arXiv Primer).

The organizers and administrators of arXiv are very explicit about the fact that they are not publishers and that they do not play the role of peer-reviewers. Instead, the moderation process is in place to ensure that relevant (topical, thematic) pre-prints and other artifact are submitted. Therefore, arXiv is a repository of at least the following materials: a) publishable materials that have been or will be published (after being peer-reviewed), b) publishable materials that are not published (not accepted for publication in the commercial journals), and c) publishable and relevant materials that may have not been submitted for publication in commercial journals (such as conference proceedings). Thus, arXiv becomes a conduit that enables knowledge artifacts to enter the disciplinary knowledge ecosystem; these are knowledge artifacts that otherwise would not have entered the disciplinary knowledge ecosystem because they would not have been published in the commercial journals.

“If you are a new user or are submitting to a new archive, you may be required to find endorsement before your submission will be processed. Users with recognized academic affiliations may be exempt from the endorsement process, while other users should contact eligible arXiv endorsers to verify that they are active members of the scientific community. This process helps restrict arXiv submissions to relevant and legitimate research contributions without adding to the administrative cost of arXiv, and thus it is an essential contribution to both the legitimacy and the sustainability of arXiv as a free resource” (arXiv Primer).

The moderation is achieved through the endorsement process, to ensure that only relevant materials make it into arXiv. Although it may seem that the endorsement process is restrictive and a form of a gatekeeper, the distributed nature of the judgment

of a particular scholar is flexible and reachable by any scholar that engages in research activities.

The arXiv Primer document covers some organizational aspects of arXiv but it is mostly a description of procedural steps about browsing, submitting, user endorsement, versioning (technical layer), and moderation (organizational layer).

“arXiv does not ask that copyright be transferred. However, we require sufficient rights to allow us to distribute submitted articles in perpetuity. In order to submit an article to arXiv, the submitter must either:

- grant arXiv.org a non-exclusive and irrevocable license to distribute the article, and certify that they have the right to grant this license,
- certify that the work is available under either the Creative Commons Attribution license, or the Creative Commons Attribution-Noncommercial-ShareAlike license, and that they have the right to grant this license, or
- certify that the work is in the public domain (we will store this information by associating the Creative Commons Public Domain Declaration with the submission)” (arXiv FAQ)

With respect to licensing issues and concerns, arXiv requests from the submitters to ensure they have the legal rights to submit their artifacts in arXiv. The intent is to enable the materials in arXiv to be accessed at no cost. Therefore, this policy in conjunction with the technical capabilities that enable the preservation of the artifacts in perpetuity is to ensure non-restricted access to materials in arXiv in the future.

#### 11.1.2 ADS (organizational)

“We are an abstracting service offering a search interface into the scientific and technical literature covering astronomy, planetary science, physics, and the arXiv e-prints. You can use our service to search the contents of hundreds of publications in these fields and then optionally access the fulltext of articles of interest to you if and when it is available online” (ADS FAQ).

“The NASA database contains abstracts from hundreds of journals, publications, colloquia, symposia, proceedings, and internal NASA reports” (ADS FAQ).

The ADS organizers define ADS as an abstracting service that provides visibility and discoverability capabilities for scholarly materials used by researchers of astronomy, planetary science, and physics. Its ability to search arXiv is specifically emphasized because of arXiv's role in the everyday life of astronomers. By enabling an enhanced accessibility to scholarly materials from many different sources in one central place, including abstracts from sources other than journals such as colloquia, symposia, proceedings, and internal NASA reports (ADS FAQ), ADS has the potential to increase the visibility and accessibility of the articles from hundreds of publications contained therein.

“The abstracts database contains data from several sources, including NASA's Scientific and Technical Information group (STI), journal publishers, SIMBAD, NED, and typed from table of contents. The original primary source of abstracts has been NASA's STI, which provided abstracts from 1975 through about the middle of 1995. Since then, we have been receiving the majority of our abstracts directly from the journal editors. Starting in 2007 we have been able to further complete our coverage of historical records thanks to our access to CrossRef, which collects and makes available metadata records for most of the content produced by Scientific journals” (ADS FAQ).

“ADS is now indexing proposals from several major institutions, including Chandra, HST, and IUE. Although we recognize that these are not published literature, they were already publicly available through the websites at those institutions, and by including them in the ADS we are allowing users quick access to the proposal data, which was otherwise difficult to get to. Successful observing proposals are thoroughly scrutinized by the telescope committees, and therefore represent significant scientific content” (ADS FAQ).

In these statements ADS is presented as a complete set of resources available to the scholarly community, having gathered scholarly materials and abstracts from multitude of journals, conferences, as well as technical reports from scientific institutions and raw data repositories of astronomical objects and observations. This comprehensive collection of resources that include abstracts of articles, technical reports, proposals, as

well as raw data repositories, position ADS as a critical actor with the ability to realign all aspects of scientists' individual knowledge production contexts. In addition, because scientists are integral actors that contribute to the disciplinary knowledge ecosystem, by positioning itself as an actor that can provide an almost complete view of the disciplinary knowledge ecosystems, ADS has the potential to impact scientists' perception of trust about ADS and arXiv.

“There should not be any significant delay in getting a paper from one of the major journals into our system, which is updated weekly. Records which we receive directly from journal publishers typically go into the system on or before the publication date” (ADS FAQ).

In addition to its comprehensiveness and inclusiveness of materials that can be accessed through ADS, there is an emphasis to the time factor and the speed with which ADS can collect the latest research findings and make them available to the scholarly community much earlier and before their publication date.

“In 1999 we started extracting reference lists from the full-text of papers available in the ADS article service or provided to us by the journal publishers. In October 2007 we gained access to CrossRef's metadata, which includes references from many articles in the physical sciences. As of June 2009 we have parsed and identified over 36 million references from all the sources of bibliographic metadata available to us” (ADS FAQ).

ADS provides a complete citation analysis capability based on the availability of CrossRef's metadata, full text articles, and other bibliographic data (ADS FAQ). The completeness of ADS and its extensive bibliographic linking capabilities have been further emphasized, providing strong incentive for its use by the scholarly community as well as by tenure committees in assessing tenure candidates' citation impact. Thus, there is an emergence of a structure (i.e., hyper-linking from article's bibliography to the articles, either abstracts or full text) that is feasible only in the digital realm. ADS is thus

implicated in scientists' knowledge production process, realigning the way scientists discover relevant materials for their research, and it can realign their knowledge networks by enabling them to fairly quickly build a picture of the research problem by traversing through articles and their bibliographies with ease.

### 11.1.3 PhilSci (organizational)

“A preprint server is used by scholars to circulate new work. A preprint is an early version of new work often in preliminary form. The archive is intended to supplement or replace an older mechanism for circulation of new work. An author used to prepare multiple copies of a new manuscript and mail it to scholars for their information and for response. Greater circulation can be achieved by posting on the archive at no cost to the author. Individual scholars can then be alerted efficiently to the preprint by informing them in a brief email of the preprint's unique ID code. Alternatively, scholars may subscribe to receive regular email updates of postings to the archive in areas of the philosophy of science of interest to them” (PhilSci FAQ).

Similarly to the intended goals of arXiv, PhilSci's administrators and organizers state that the intended role of PhilSci is to provide a mechanism for the exchange of the current works of scholars in a preliminary form (“early version of new work”) in order to receive feedback (“for their information and for response”) from other scholars, supplementing the traditional paper based manuscript circulation process among the community of philosophers of science. One of the intended outcomes and values that PhilSci can provide is to enhance circulation of scholarly work. Thus, PhilSci has implications for scholars' knowledge production process by enabling them to exchange early versions of their manuscripts, some of which might be work in progress such as conference papers, using electronic means, and also reconfiguring the body of knowledge artifacts that are available for scholarly use—the included early versions of

manuscripts and conference papers can be searched, discovered and accessed by scholars elsewhere.

“A journal publishes material that has passed scrutiny by referees and has been edited by the editorial staff to bring it to the journal standards. The archive does not referee postings and does not edit them. The archive merely filters minimally to assure relevance to philosophy of science” (PhilSci FAQ).

“Preprints posted on the archive are restricted to those in philosophy of science or related material of interest to professional philosophers of science. The range of admissible topics and the style of analysis is set by the topics and styles of material publishable in the Philosophy of Science Association journal, *Philosophy of Science*” (PhilSci Policy).

Similarly to arXiv, PhilSci organizers caution users of PhilSci that it is not a publishing venue and that its role is to ensure that the submitted materials are relevant to the discipline, without any editing of the submitted materials, and without passing any judgment on their quality or scholarly merit beyond disciplinary relevance.

“The archive does not require or expect that material has not appeared elsewhere. However, if it has appeared elsewhere, the author must determine whether copyright was transferred from the author and whether the copyright agreement allows posting on the archive. While we do not object to duplication, the other source may” (PhilSci).

PhilSci organizers also make a point to notify the submitters of materials that they should ensure with publishers (or other entities that have the copyright to previously published materials) that they have the right to publish in open access archives. Thus, through Copyright Transfer Agreements (CTA), as the documents that enact the regime of how certain copyrighted work can be distributed and used, the publishers (as copyright holders) have the potential to effect scholars’ decisions about what to be submitted in PhilSci.

“These categories are dedicated to particular conferences and volumes. Only contributors designated by the conference organizers or volume editors may post papers” (PhilSci FAQ).



PhilSci is intended to act as an exchange mechanism to enable conference attendees to share the conference submitted papers ahead of time in preparation for a specific conference (as a conference proceeding alternative), realigning scholars' knowledge production process to consider conference papers before the conference takes place, moving to the pre-print model of the sciences. This in turn enables the conference papers and proceedings to enter the disciplinary knowledge network by making them visible and accessible beyond the immediate time and space of conference event.

“The PhilSci Archive is offered by its sponsors as a free service to philosophers of science. Its goal is to preserve and foster the rapid exchange of new work in philosophy of science” (PhilSci Policy).

Again, similarly to arXiv's goal to enhance and accelerate the research in the specific field such as astronomy and astrophysics, PhilSci is also intended to foster and enable an accelerated exchange of scholarly works for philosophers of science.

#### 11.1.4 Protocols and standards

“The roots of OAI lie in the development of e-print repositories (so-called archives). E-print repositories were established in order to communicate the results of ongoing scholarly research prior to peer review and journal publication” (History and development of OAI-PMH).

The initiators and developers of the OAI fully acknowledge the co-constructionist nature of OAI related standards and protocols, with respect to enabling scholarly work to be distributed widely and quickly via electronic archives before their official publication.

“The Open Archives Initiative [(OAI)] develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The Open Archives Initiative has its roots in an effort to enhance access to e-print archives as a means of increasing the availability of scholarly communication. Continued support of this work remains a cornerstone of the Open Archives program. The

fundamental technological framework and standards that are developing to support this work are, however, independent of the both the type of content offered and the economic mechanisms surrounding that content, and promise to have much broader relevance in opening up access to a range of digital materials. As a result, the Open Archives Initiative is currently an organization and an effort explicitly in transition, and is committed to exploring and enabling this new and broader range of applications. As we gain greater knowledge of the scope of applicability of the underlying technology and standards being developed, and begin to understand the structure and culture of the various adopter communities, we expect that we will have to make continued evolutionary changes to both the mission and organization of the Open Archives Initiative” (About OAI).

The intended goal of the OAI interoperability standards is to provide the scholarly community with tools for easy, fast and efficient exchange of scholarly works. Although OAI is neutral to the type of content and scholarly communities that utilize the standards, they are aware that cultures within different scholarly communities vary. In the spirit of openness, the OAI organization positions itself as a structure that is modifiable because of its role in the scholarly community context. Here, the co-constructionist approach of OAI is explicitly stated and recognized that it is work in progress. The development of OAI protocols and standards is contextualized within the information practices and disciplinary cultures. Thus, OAI intentionally learns from different scholarly communities about the features and functionalities it needs to build and make available for them in the future.

“As an organisation, the OAI has included an Executive for management, and Steering and Technical Committees for policy direction and evaluation of protocol developments. The Digital Library Federation (DLF), the Coalition for Networked Information (CNI), and the National Science Foundation (NSF) have funded the OAI. While the Executive and the funders are USA-based, the success of the OAI is firmly grounded in the participation of a community of people from around the world, particularly Europe as well as North America. Now that there is a well-developed and stable second version of the protocol, the need to keep control in the hands of a very small number of people who can take independent and speedy decisions may be less important when weighed against the perception of stability and authority conferred by control through a standards

body such as ISO, and this possibility has been discussed within the OAI” (OAI For Beginners).

“Policy decisions about the Open Archives Initiative are made by a Steering Committee. The interoperability infrastructure was developed by a technical committee, which continues to advise on the infrastructure as experience with it develops. Herbert Van de Sompel and Carl Lagoze are responsible for coordination of OAI activities, which are centered at Cornell University” (OAI FAQ).

To enable broader reach of its protocols with intent to standardize them, OAI’s management has incorporated in its management important institutional actors, such as the Digital Library Federation (DLF), the Coalition for Networked Information (CNI), and the National Science Foundation (NSF), that either fund OAI’s work or are involved in other decision making aspects. The inclusion of these institutional actors in the funding and development of OAI projects puts them in a position to realign scholarly production process and knowledge networks—using OAI as the intermediary. In turn, the availability of open standards and protocols means that they can be used by different scholarly communities to establish open access repositories or build access tools based on open standards.

The significance of the organizational and the technological layer is analyzed in the following section.

### *11.2 The technological perspective*

In this section the technological capabilities, features and available functions of the access tool, open access repositories and the related protocols and standards are described as articulated by the organizing and coordinating structures that built them. They are further analyzed and interpreted with respect to their potential ability to realign various elements that constitute researchers’ individual knowledge production contexts.

The evidence is extracted from the documents that describe the technological features and capabilities.

As it was explained earlier at the beginning of section describing the organizational perspective, the distinction between the technological and organizational layer of the open access repositories and access tools is meant to delineate between the technical level such as “EPrints”, and organizational level such as the PhilSci archive. So, the PhilSci archive is constructed by using a technology (EPrints) that is used to build a specific disciplinary repository (PhilSci). Understanding this distinction is important as it will help with the understanding about how technology is implicated in the construction of knowledge production contexts. As it will be shown later, the distinction between the technological and organizational aspect of arXiv (and ADS) are not as clearly and distinctly delineated in the documentation, because both arXiv and ADS have been constructed as “home built” systems within the disciplinary boundaries.

The technical layers of the OA repositories and their corresponding access tools are analyzed and interpreted first (arXiv, ADS, PhilSci, EPrints, Google Scholar), as well as the protocols and standards used by the repositories and the access tools. This analysis is based on the documentary evidence.

### 11.2.1 arXiv (technological)

“All arXiv submissions are freely available, often in multiple formats. Each submission has an "abstract" page where summary metadata for the submission are displayed, along with the available download formats and any additional services that apply to that particular submission. Examples of extra services include trackbacks, SLAC-SPIRES reference data, CiteBase citation information, and chronological or subject-specific arXiv browsing via "previous" and "next" links. Please note that some of the services mentioned are third-party services that may not be available for every submission and that are maintained by

different administrative staff. Check carefully any of the offered pages for additional information and separate help documentation” (arXiv Primer).

The technical capabilities for organizing the submitted artifacts into arXiv are described, emphasizing the capabilities that enable arXiv artifacts to be easily discovered and accessed, either directly via its own portal or via extra services available outside of the arXiv structure (such as ADS that is used to access arXiv for resources related to astronomy and astrophysics). The intent of these technical capabilities is to enable arXiv’s integration with external services and thus enable its artifacts to be easily discovered and accessed, either by researchers or by other systems. Thus, these technical capabilities increase the visibility of the materials deposited therein.

“arXiv submissions are meant to be available in perpetuity. Thus, arXiv has high technical standards for the files that are submitted. The submission process begins with the preparation of valid metadata for the paper and continues through a verification process whereby the files uploaded are checked for certain problems. Valid upload formats for the primary text file include (La)TeX, PostScript, PDF, and HTML. There are specific restrictions for each of the primary file types and only certain other supplementary file types are acceptable; please consult the submission instructions for additional information. If there are problems with the metadata, error messages on the screen should help you make corrections. If there are problems with the files, an error report will be generated and the files will be assigned a rejection id. In this case you should read the log and attempt to correct the problem, or failing that send the rejection id to the arXiv administrators for assistance” (arXiv Primer).

The capabilities of arXiv have been built with the intent to enable long-term availability of the artifacts submitted therein. It is noted that long-term availability can be achieved by preparing the form and format of articles to maximize their accessibility, as well as structure arXiv in such a way to allow for different formats and file standards to be submitted and accessed. arXiv’s technical capabilities are described in more details in the arXiv FAQ, primarily a technical document covering topics such as: a)

downloading and viewing files (formats, MIME types, linking services such as OpenURL and RSS feeds to the entire archive or specific sections, interface standards such as OAI, establishing mirror sites), b) procedures, formats, and steps for formatting and preparing submissions to arXiv, licensing considerations, etc., and c) contextualizing of arXiv materials (references to and from arXiv document, trackback and links to blogs, and social bookmarking ability). Thus, arXiv's intended technical ability to make the materials deposited therein available in perpetuity can realign researchers' knowledge network (by enabling access to old materials and the knowledge contained therein) and it can realign the research process by making old materials centrally available and easy to find and access.

“arXiv supports and participates in the Open Archives Initiative (OAI). arXiv is a registered OAI-PMH data-provider and provides metadata for all submissions which is updated each night shortly after new submissions are announced” (arXiv FAQ).

Enhancing the visibility of the materials is achieved via its participation in the OAI that enables other external services to harvest and download the article level metadata from arXiv on a daily basis. This is the primary way through which ADS harvests arXiv daily and it is able to provide a comprehensive and central approach for searching arXiv submissions in addition to other sources. Thus, by supporting and participating in OAI, arXiv can enhance article's visibility that in turn makes it easier for scientists to search, discover and access the knowledge artifacts needed for their research.

To ensure maximum use of the content that is managed, the arXiv FAQ is concerned even with the most minute of details, such as the line level formatting of abstracts:

“Abstracts are automatically processed, so it is important to adhere to the format. The most important processing is line-wrapping to 80 characters. This is preferable for most abstracts, but there are occasions when lines shouldn't be

wrapped together, e.g. if the abstract is a table of contents. To allow for this, lines with leading white space are not wrapped. Thus, you should not submit an abstract that contains leading white space unless you explicitly prefer no wrapping” (arXiv FAQ).

Again, there is an emphasis on making sure that metadata for each article is carefully structured to ensure proper processing by arXiv for further use in its own portal or by external services that add to increasing the visibility, discoverability and accessibility of the materials deposited therein.

### 11.2.2 ADS (technological)

As an access tool, ADS provides search capabilities not only for arXiv, but also for two additional bibliographic databases. This point is further described in the ADS User Guide:

“... [ADS is] a powerful search engine for each of its three bibliographic databases:

- Astronomy and Astrophysics (including abstracts from Planetary Sciences and Solar Physics journals)
- Physics and Geophysics (including abstracts from APS journals and SPIE proceedings)
- ArXiv e-prints (all the papers published in the ArXiv e-print archive)”

Further,

“The databases cover all the major journals, many minor journals, conference proceedings, many Observatory reports and newsletters, many NASA reports, and PhD theses” (ADS User Guide).

ADS has thus been constructed with a set of features and functional capabilities enabling it to be built as a very complete and comprehensive search engine covering multitude of sources (raw data, astronomic objects, abstracts, etc.) and artifacts formats (such as PDF, LaTeX, links to full texts, etc) as listed in the ADS User Guide.

“The ArXiv e-print database contains preprints submitted to the arXiv e-print archive. The ADS maintains this database to allow searches on the latest

literature being published, with links to the full text available from the ArXiv” (ADS User Guide).

“We provide access to scanned images of articles from most of the major and most smaller astronomical journals, as well as many conference proceedings series. All scanned articles are linked to the corresponding references in the ADS” (ADS User Guide).

“As of March 2005, references the from arXiv preprints are integrated in the ADS. When we retrieve the metadata for the nightly update of arXiv preprints, we also process the source data to retrieve the references contained in it, either from the (La)TeX source or the PDF version of the preprints. Next, the retrieved references are parsed to resolve them into a match with an existing record in the ADS” (ADS Help).

Specific mention is made of arXiv, with ADS enabling arXiv to be searched in a more comprehensive way via a tool that can enable serendipitous possibilities because many different sources are being brought together and integrated for searching purposes. As it has been already noted, ADS also provides access to literature (scanned old journal articles) not available via the commercial journals. Therefore, ADS not only has the ability to make it easier for scientists to find the knowledge artifacts they need for their work (process realignment), it also realigns scientists’ knowledge networks by including and making visible literature that otherwise might not be accessible due to its age and non-availability in digital form.

“Users can query the database by author, astronomical object name/position (astronomy database only), words in the title, and words in the abstract text. The "results list" is ranked by how closely the paper matches the query (unless otherwise requested by the user). From the results list one can view the full record available for each of the returned articles, including scans, HTML and PDF versions of articles, if they are available” (ADS User Guide).

“We provide references and citations whenever possible. The references and citations database is excellent, but, as is the case with all such databases, it is not complete” (ADS User Guide).



In addition to its system level capabilities that enable ADS to support multitude of submission types, file formats, and linking capabilities, ADS provides advanced search capabilities to scientists with the intent to enable them to discover easily the latest research findings and other artifacts for their knowledge production. The search interface is very powerful and flexible with many selection and narrowing criteria for queries (ADS Guide). In addition to searching, ADS can be browsed by journal title, table of content, volume, year, etc. (ADS User Guide).

The output of the search query can be formatted (per scientists' feature selection) so that they can be imported automatically into reference and bibliographic tools such as EndNote, BibTeX, ProCite, RefMan, RefWorks, Dublin Core XML, and others (ADS Help).

ADS also provides a "Find Similar Abstracts" feature that might be useful for scientists to identify similar corpus of articles in order to identify patterns of research activities around a specific topic of interest (ADS Help).

"With about 75% of all refereed astronomy papers appearing first in the arXiv, and well over 90% of the highly cited ones (99 out of the top 100 2003 ApJ papers were preprinted, according to our recent study), the preprints have become an integral part of astronomy (and physics) research" (ADS Help).

"The use of the preprints has a major, positive effect on research. The effective latency between publication and citation has shortened by about six months, without taking the citations to the preprints themselves into account. This has the effect of increasing the rate of discovery; we believe we have an obligation to support this change" (ADS Help).

Here ADS provides a justification, based on citation impact related research studies, for the inclusion of the entire corpus of arXiv materials into its search engine. The main argument that is provided is that arXiv has enabled an increased rate of discovery by increasing article visibility, discoverability and accessibility. As a result, ADS has

positioned itself and its capabilities to further enhance arXiv in its mission, by realigning scientists' knowledge production processes and knowledge networks. This is an example of how technological level features have been co-constructed between arXiv and ADS in the discourse of the managing structures and by the structuring of the technology, performing on each other to further strengthen the value that emerged from their ongoing integration.

### 11.2.3 PhilSci (technological)

From technological perspective, PhilSci runs on the open source software available from eprints.org: "This archive is running on eprints.org open archive software, a freely distributable archive system available from eprints.org" (About PhilSci). The specific version of EPrints that PhilSci runs on is version 2.2.1, published in 2002 (the most current version is 3.1.3, published in May 2009). Thus, for the purpose of this analysis the features and functional capabilities of EPrints version 2.2.1 are described.

The PhilSci Help documentation that describes the technical capabilities of PhilSci is mostly procedural and a how-to guide explaining how the interface can be used by scholars for searching the archive as well as deposit material therein. Topics covered in the PhilSci Help document are: a) Browsing approaches, b) Searching approaches, c) Registering to deposit in PhilSci, d) Depositing procedures, e) The user workspace within PhilSci and the features available, f) Deposit types, g) Bibliographic information, h) Subject categories, and i) Document submission formats.

Similarly to arXiv, PhilSci also supports a multitude of features and functional capabilities related to artifact types, supported file formats, search capabilities and linking capabilities with the intent to enable increased artifact visibility, discoverability,

and accessibility. As it is described in the next sub-section, many of the technical capabilities of PhilSci derive from the capabilities and features provided by the EPrints software.

#### 11.2.4 EPrints

The EPrints software version 2.2.1 runs on the LAMP open source software stack (EPrints 2.2 Documentation, 2002, p. 9-13). LAMP is a label denoting a combination of different open source software that when combined together enable running EPrints and other tools free of costs, due to their compliance with various open source software licenses. LAMP is defined as: Linux (operating system) + Apache (web server) + MySQL (database) + Perl (scripting and programming language).

“GNU EPrints is generic archive software under development by the University of Southampton. It is intended to create a highly configurable web-based archive” (EPrints 2.2 Documentation, 2002, p. 7).

“GNU EPrints primary goal is to be set up as an open archive for research papers, and the default configuration reflects this, but it could be easily used for other things such as images, research data, audio archives - anything that can be stored digitally, but you’ll have make more changes to the configuration” (EPrints 2.2 Documentation, 2002, p. 7).

It is important to note that the EPrints software is independent to its use as enabler of the PhilSci repository, and it is neutral to its use for pre-prints. It can also be used for other types of digital objects such as images and audio files. The instantiation of EPrints into a PhilSci is one of the many ways EPrints can be used. As of November 30, 2009, ROAR lists 353 repositories built using EPrints. For example, The Linnean Connections (<http://www.linnean-online.org/>) is a repository of images of plants, fish, shells, and insects from the The Linnean Society of London.

“The system has been designed to encourage better quality data - that is to say accurate and unambiguous. Well defined metadata can really help if you want to start exporting your archive data or making it interoperable with other systems.” (EPrints 2.2 Documentation, 2002, p. 7)

“Once you register your archive (at <http://www.openarchives.org>) various search systems will be able to collect the metadata (titles, authors, abstract etc.) and allow more people to find records in your archive” (EPrints 2.2 Documentation, 2002, p. 18).

Similarly to arXiv, EPrints emphasizes on the features and capabilities to enable open archives by making it easy for its content to be shared and incorporated into external systems. Thus, EPrints 2.2.1 has full support for OAI version 1 and version 2 (EPrints 2.2 Documentation, 2002, p. 18), enabling its managed content to be shared with any OAI compliant systems over the OAI-PMH interface.

EPrints enables PhilSci’s browsing and search capability from its own portal. It also provides a capability for subscription updates to be sent via e-mail on regular basis to users that have subscribed for updates based on keywords and other search criteria (EPrints 2.2 Documentation, 2002, p. 18-19).

#### 11.2.5 Google Scholar

“Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. Google Scholar helps you identify the most relevant research across the world of scholarly research” (About Google Scholar).

Unlike ADS that is a specialized search engine for arXiv, PhilSci does not have its own specialized search engine. Instead, the PhilSci scholars stated during the interviews that they use PhilSci’s own portal, Google, or Google Scholar when searching for

PhilSci materials. Google Scholar is incorporated in this analysis as an example of non-disciplinary or non-thematic access tool.

According to the About Google Scholar document, Google Scholar makes possible to: a) search diverse sources from one convenient place, b) find papers, abstracts and citations, c) locate the complete paper through your library or on the web, and d) learn about key papers in any area of research.

“Google Scholar can boost the worldwide visibility of your content. We work with scholarly publishers to index works from all research disciplines and make them searchable on Google Scholar. Learn more about our policies and find technical information for scholarly publishers and societies” (About Google Scholar).

The expressed and intended goals of Google Scholar as an access tool are similar to those expressed by ADS. Although, unlike Google Scholar, ADS is a specialized tool and can offer features and capabilities that integrate it with arXiv in a more contextual and coherent way. For example, due to its uniform metadata standards (based on OAI) and interface protocol for metadata exchange (OAI-PMH), ADS also provides an easy approach for browsing and traversing between articles by clicking—by the way of clicks into article’s references to access the full texts that also includes its own clickable bibliography.

Google Scholar does not support OAI at this time. Google Scholar’s support for OAI ended in 2008 according to Mueller (2008). Instead of OAI it utilizes the Google crawling engine, using a more generic XML based interface, the XML Sitemap.

#### 11.2.6 Protocols and standards

“Mapping among multiple metadata formats would place a considerable burden on service providers, who harvest the metadata and use it to build higher level services. While there is research work on creating services such as common

search interfaces across heterogeneous metadata formats, a less burdensome and ultimately more deployable solution is to require repositories to map to a simple and common metadata format. The fifteen elements Dublin Core has over the past several years evolved as a de facto standard for simple cross-discipline metadata and is thus the appropriate choice for a common metadata set. Cooperation between the OAI and the Dublin Core Metadata Initiative has led to a common xml schema for unqualified dublin core” (OAI FAQ).

The increased article visibility, discoverability and accessibility for arXiv and ADS are achieved via a standardized approach for meta-data description and meta-data exchange interfaces. In the context of scholarly exchange, the Dublin Core (DC) metadata elements have been appropriate by software developers and archive providers in conjunction with OAI. This provides for simpler, standardized, and normalized mechanism for the representation and exchange of metadata describing scholarly artifacts.

“OAI-PMH is a low-cost mechanism for harvesting metadata records from one system to another – from Data Providers to Service Providers. Multiple Service Providers can harvest from multiple Data Providers ensuring a wider spread of metadata. OAI-PMH is not a search protocol, but its use can underpin search-based services; it is a base layer on which to build other services” (History and development of OAI-PMH).

The OAI-PMH is a standardize protocol that enables specialized services to harvest artifact metadata on regular basis. ADS is one such service that harvests arXiv’s metadata on regular basis using OAI-PMH. From the perspective of OAI-PMH, PhilSci and arXiv are data providers (OAI FAQ) and ADS is a service provider: “A service provider issues OAI-PMH requests to data providers and uses the metadata as a basis for building value-added services” (OAI FAQ).

Considering that organizational structures and technological capabilities are intertwined and mutually inform each other as it has been already shown, the nature of

the co-construction dynamics are analyzed next, contrasting and comparing the OA repositories access tools used by the two groups of participants.

### *11.3 The symbiotic relationships between the organizational and the technological layers*

Both arXiv and PhilSci's organizers, managers, sponsors and administrators are driven by the same goals to enable their respective scholarly communities to rapidly access the latest research findings needed for their knowledge production, as emerged from the documentary evidence and it will be analyzed and interpreted in more details in the rest of this section. To achieve these goals, numerous technological features and functional capabilities from the technological layer together with the enactment of policies that manage the repositories, have been inscribed (through adaptation and appropriation) into arXiv and PhilSci to structure them as socio-technological information environments that can enable long term accessibility to scholarly artifacts. "Appropriation" means that features and functional capabilities from the technological layer are actually implemented and can be used by the scholars. For example, not all of the features made available by the EPrints software are necessarily implemented for use in PhilSci. At other times, the technological features and capabilities are "adapted" in the process of being implemented for use. The adaptation process is sometimes as simple as changing configuration setting, and at other times might mean additional development. As it has been shown in the previous two sections via statements articulated by the organizers of PhilSci, arXiv and ADS, in the mission statements of the respective documents, and by the developers of the technological layers of the repositories and the

access tools, the technological and the organizational layers have constantly performed on each other by inscribing features and capabilities onto each others' layers. For example, the development of open standards and protocols (versus proprietary standards and protocols) by the OAI which are specifically designed to describe and exchange scholarly artifacts, have been quickly appropriated (i.e., inscribed) and used in software (such as EPrints) that can enable the establishment of an open access repository such as PhilSci. The documentary evidence reveals symbiotic and intertwined relationships amongst the technological layers of the standards and protocol, the access tools software and the software used to build the repositories, and between the technological and organizational layers. These are described in the rest of this section.

Based on the documentary evidence presented in sections 11.1 and 11.2, the following four tables, Table 22, Table 23, Table 24, and Table 25 present summary views of the technological and organizational layers of the open access repositories and access tool.

Table 22: Comparison of the technological layers of features and functional capabilities of arXiv and PhilSci

Available features and functional capabilities	Supported and enabled by the software used to build arXiv and PhilSci	
	Home built (for arXiv)	EPrints v2.2.1 (for PhilSci)
Multiple file formats	Yes	Yes
Linking	Yes	Limited
Searching	Yes	Yes
Browsing	Yes	Yes
Metadata exchange	Yes	Yes
Different types of artifacts	Yes	Yes
Daily updates	Yes	Yes
OAI	Yes	Yes
OAI-PMH	Yes	Yes
Sub-categorization	Yes	Yes

Note: The technological layer is comprised of the features and functional capabilities made available by the software components that are used to build the repository.

Sources: Based on document analysis.



Table 22 is constructed by identifying the features and functional capabilities of the technological layers related to arXiv and PhilSci. The features and functional capabilities of both arXiv (home built software) and PhilSci (EPrints) are aggregated in column 1. Column 2 and 3 respectively denote whether the repositories provide the specific feature or capability. For the most part a same set of technological features and capabilities have been inscribed into the open access repositories. Based on the documentary evidence, the software used to build arXiv and PhilSci has been performed upon (by actors such as the standards and protocols that are both global and local actors; they are global as “standards” that are locally implemented for a specific technology) to enact the following capabilities: support multiple file formats (.doc, .pdf, etc.), searching via different types of queries (keyword, authors, date, etc.), browsing, meta-data exchange (specifically support for OAI and OAI-PMH), support for different types of artifacts (articles, conferences submissions, reports, etc), providing updates by e-mail subscriptions, as well as sub-categorization of the submitted artifacts into different categories.

A main point of differentiation is that EPrints v2.2.1 that is used to build PhilSci does not support any systematic and automatic linking between abstracts and full text of articles or pre-prints, nor does it support any bibliographic linking between artifacts. Further, the EPrints software is designed and developed independently of PhilSci and it powers many more open access repositories. The inscription process in how EPrints acquires its technological features and capabilities is independent of the philosophy of science discipline (i.e., EPrints is not specifically built with PhilSci in mind, or with the needs of the philosophers of science) and there is clear delineation between EPrints as

the technological layer and PhilSci as the organizational layer. EPrints is developed as an open sources software at the School of Electronics and Computer Science, University of Southampton, UK, that is appropriated by the Department of History and Philosophy of Science, University of Pittsburgh to establish PhilSci as an open access repository for a specific discipline. In arXiv however, the inscription process of the technological layer is not as clearly delineated from the inscription process of the organizational layer—except in the way elements from the technological and the organizational layers are articulated in the documentary evidence that describe arXiv. Both the technological and the organizational layers of arXiv are developed, built, supported and maintained by the same entity, the Cornell University Library, thus the technological layer of arXiv is not necessarily articulated separately from the organizational layer in the documentary evidence related to arXiv.

Table 23: Comparison of the organizational layers of intended roles and values of arXiv and PhilSci

	<b>Intended by the organizing and management structures of arXiv and PhilSci</b>	
<b>Intended roles and values</b>	<b>arXiv repository</b>	<b>PhilSci repository</b>
Open access	Yes	Yes
Central location	Yes	Yes
Supplement the traditional scholarly process	Yes	Yes
Enhance researchers search process	yes	Yes
Search portal	Yes	Yes
Browsing portal	Yes	Yes
Relevance submission control	Yes	Yes
Conference papers accepted	Yes	Yes
Different types of artifacts (articles, conference submissions, technical papers, etc.)	Yes	Yes
Materials available in perpetuity	Yes	No (materials can be removed)

---

Note: The organizational layer is comprised of the intended roles and values that the repositories can provide the respective researchers with. These are roles and values as intended by the organizing and management structures of arXiv and PhilSci, achieved by inscribing a combination of the available features and functional capabilities from the technological layer, and policies.

Sources: Based on document analysis.

Table 23 is constructed by identifying the intended roles and values that arXiv and PhilSci ought to provide for the respective disciplines. The intended values and roles are articulated in the documentary evidence that describe the organizational layers of arXiv and PhilSci. The intended roles and values for both arXiv and PhilSci are aggregated in column 1. Column 2 and 3 denote whether the respective repositories describe themselves as enablers of these roles and values for their scholarly community. The intended roles and values are partially inscribed into arXiv and PhilSci by appropriating features and capabilities from the technological layers, but more importantly, they are a result of an inscription of purposefully structured intents, goals and policies into PhilSci and arXiv, intended to help researchers in their knowledge production and enhance the scholarly exchanges in the specific discipline. Both arXiv and PhilSci are intended to enable open access to scholarly works, become a central location for exchange of the latest research findings, supplement the traditional scholarly process by enabling sharing of pre-prints much earlier and before their publication in commercial journals, and enable discoverability of the articles via searching and browsing. They both intend to support different types of artifacts, including conference submissions, via a relevance submission control mechanism to ensure disciplinary relevance of the materials that are submitted. There is one notable difference that has emerged from the technological and organizational analysis. While arXiv intends to keep the submitted artifacts in perpetuity (based on enacted policy), PhilSci has a mechanism (and enacted policy) by which submitted materials can be removed from PhilSci. Thus, as it is summarized in Table 22

and Table 23, from the technological and organizational perspectives, both PhilSci and arXiv provide almost the same set of feature and functional capabilities that are intended to provide the same set of goals and values to their respective scholarly communities.

Table 24: Comparison of the technological layers of features and functional capabilities of the access tools used by arXiv and PhilSci participants

Features and functional capabilities	Supported and enabled by the software used to build ADS and Google Scholar	
	arXiv used via ADS	PhilSci used via Google Scholar
Specialized	Yes	No / generic
Multiple sources	Yes	Not defined
Artifacts (articles, reports, pre-prints, etc.)	Yes	Yes
Abstracts	Yes	Not clearly defined
Interoperability with repositories	Yes	Yes, no standard approach
Bibliographic links (hyperlinks within and outside)	Yes	Yes, non consistent
OAI	Yes	No
OAI-PMH	Yes	No
Search capability	Yes	Yes
Brows capability	Yes	No

Note: The technological layer is comprised of the features and functional capabilities made available by the software that is used to build the repository.

Sources: Based on document analysis.

Table 24 is constructed by identifying the features and functional capabilities of the technological layers of ADS and Google Scholar. The purpose of this table is to compare the technological features and capabilities between ADS and Google Scholar that will inform the analysis when comparing the organizational layers of ADS and Google Scholar. Both ADS (used by astronomers) and Google Scholar (used by philosophers of science) emerged as the respective access tools based on the interviews with the researchers. The features and functional capabilities of ADS and Google Scholar aggregated in column 1, are based on the articulation as described in the documents that

describe ADS and Google Scholar. Column 2 and 3 respectively denote whether the repositories provide the specific feature or functional capability. Unlike the technical layers of arXiv and PhilSci, where they differ only with respect to the linking capability they provide, the access tools have been inscribed with features and functional capabilities that vary greatly. While ADS has emerged from the inscription process as a specialized access tool with a clear intent to build features and capabilities that may enhance the scholarly information practices of astronomers, philosophers of science do not have a specialized search engine. Rather, the PhilSci participants stated during the interviews that besides the PhilSci portal, Home Pages of scholars, and invisible colleges, they use Google Scholars as an access tool to discover and access materials in PhilSci. Google Scholar, given that it is not a specialized access tool for PhilSci, does not necessarily harvest all of the sources that may benefit philosophers of science, nor does it clearly define its harvesting of abstracts related to the philosophy of science discipline. Since it dropped the support for OAI-PMH in 2008, it is not clear whether Google Scholar has implemented a systematic way to harvest scholarly artifacts from PhilSci. Google Scholar still provides search capability and PhilSci's materials can be discovered through it by searching; however, there are limited bibliographic linking capabilities. Also, Google Scholar does not provide browsing capabilities for PhilSci materials—Google Scholar does not have topical or disciplinary categorization of its indexed content for browsing purposes.

ADS on the other side is strongly integrated with arXiv and has a stated goal to harvest all possible materials (pre-prints, post-prints, scan older journals, abstracts, raw data catalogs, etc.) related to astronomy and astrophysics. The interoperability and

metadata exchange capability is achieved by supporting the latest OAI and OAI-PMH standards and protocols. Further, ADS provides a value add service where they proactively build linking capabilities between the different types of materials (such as abstracts, pre-prints, post-prints) and also provide a powerful bibliographic linking capability that may perform on and enhance scientists' research processes.

Table 25: Comparison of the organizational layers of intended roles and values of the access tools used by arXiv and PhilSci participants

Intended roles and values	Intended by the organizing structures for ADS and Google Scholar	
	ADS	Google Scholar
Search interfaces	Yes	Yes
Enhances scholarly process	Yes	Yes
Enabled quick and fast access	Yes	Yes
Search multiple sources	Yes	Yes / can't be determined
Completeness in abstracting	Yes	can't be determined
Used by tenure committees	Yes	No

Note: The organizational layer is comprised of the intended roles and values that the repositories can provide the respective researchers with. These are roles and values as intended by the organizing structures of ADS and Google Scholar, achieved by appropriating a combination of the available features and functional capabilities from the technological layer.

Sources: Based on document analysis.

The specialized approach of ADS and the generic approach of Google Scholar at the technological layers are being inscribed (through adaptation and appropriation) into the respective organizational layers mediated by the organizing and management structures that have specific intentions and goals, as summarized in Table 25.

Table 25 is constructed by identifying the intended roles and values that ADS and Google Scholar ought to provide for the respective disciplines. The intended values and roles are articulated in the documentary evidence that describe the organizational layers of ADS and Google Scholar. The intended roles and values for both ADS and Google Scholar are aggregated in column 1. Column 2 and 3 denote whether the respective

access tool describe themselves as enablers of these roles and values for the scholarly community. Both ADS's and Google Scholar's intention is to provide researchers with search capabilities that increases scholarly artifacts' visibility, discoverability and accessibility, with the goal to enhance researchers' research processes. Enabling quick and fast access to the latest research findings is emphasized as a goal by ADS and Google Scholars as well; they both intend to enable researchers to search multiple sources of knowledge artifacts. While ADS has clearly shown (based on the documentary evidence) that they indeed provide scientists with the capability to search across a multitude of sources and artifact types, Google Scholar's intent cannot be ascertained due to lack of specific description in the documentary evidence or which philosophy of science sources they harvest and index for searching.

The differences in the construction process of the technological and organizational layers of arXiv and PhilSci, as well as the respective access tool used by the two communities of researchers (ADS and Google Scholar), provide an insight about the emergence of arXiv and ADS with strong performative capabilities amongst the astronomers and astrophysics and the weak performative capabilities of PhilSci amongst the philosophers of science. The inscription processes that perform on the construction of arXiv and ADS, both at the technological and the organizational layers are instantiated within the disciplinary context. In comparison, only the inscription processes that perform on the construction of PhilSci's organizational layer are instantiated within the disciplinary context of philosophy of science. The inscription processes that perform on EPrints, as the technological layer of PhilSci, and Google Scholar are instantiated outside of the disciplinary culture (by organizations that do not

build software and tools to be used specifically by philosophers of science). Therefore, the disciplinary context of astronomy and astrophysics has impacted the technological and organizational layers of arXiv and ADS, while the disciplinary context of the philosophers of science has impacted only the organizational layer of PhilSci. Therefore, for arXiv, a process of ongoing co-construction emerges where the disciplinary culture and arXiv/ADS (both technological and organizational layers) mutually inform each other, resulting in increased visibility, discoverability and accessibility of articles. For PhilSci, the process of co-construction within the discipline is limited to PhilSci's organizational layer, evident in the use of old EPrints software that was released in 2002, even though there is new and enhanced version of EPrints available that was released in May 2009. The weak performative process between the disciplinary culture and the technological layer of PhilSci (i.e., EPrints) and Google Scholar, is also evident by the use of Home Pages of scholars, invisible colleges and JSTOR, by the PhilSci scholars.

Do the above identified differences amongst and between the technological and organizational layers of access tools and repositories have implications for scholars' individual and disciplinary knowledge production contexts as perceived through the four themes? This question, and more broadly the Research Questions are explicated in Chapter 12.

#### *11.4 The impact of the technological and organizational layers on articles' visibility, discoverability and accessibility*

The co-construction dynamics within and between the technological and organizational layers of arXiv, ADS, and PhilSci are instantiated and perform on the premise that there is open content (i.e., knowledge artifacts mostly in the form of pre-



prints and post-prints) that can be made available to researchers for use in their knowledge production process. The aim of arXiv, ADS and PhilSci is thus to enable increased use of the knowledge artifacts deposited therein. The technological layers of the open access repositories and the access tools have thus over time been inscribed with features and capabilities to support different file formats, different types of knowledge artifacts, support for open standards and protocols, integration and interoperability with other system (abstracting, indexing), ability to include files from multiple sources, as well as bibliographic linking amongst individual articles, abstracts and indices. These features are congruent with the operational definition of visibility that has been adapted for this study (see section 2.8) based on the review of literature. The socio-technological co-construction between the organizational and technological layers has been maintained and strengthened over time, more so for astronomers than for the philosophers of science, and in conjunction with the increased visibility of the materials contained in the open access repositories, the discoverability (artifact's location can be determined; the artifact might or might not be accessible) and accessibility (artifact's content can be read by humans or machines) of the materials has also been enhanced, resulting in increased use, as it has been perceived by the researchers.

The co-construction dynamics between the lived experiences and the technological and organizational layer is described in more detail in the next chapter, by answering the Research Questions, elaborating on the implication of the open access repositories on scholars' knowledge production processes and the implication on the disciplinary knowledge production context.

## **Chapter 12. Study findings: contextualizing the individual and the disciplinary information practices**

In this chapter, the lived experiences and perceptions layer and the technological and the organizational layers as analyzed and interpreted in Chapter 8 through Chapter 11 are contextualized within the respective disciplinary knowledge production contexts represented by the PhilSci scholars and arXiv scientists. The relationship dynamics between and amongst the different layers are identified and categorized based on the perceived, intended and supported properties alongside their dimensions with respect to the four themes. This chapter proceeds by answering the Research Questions and continues to explicate the co-construction dynamics of individual and disciplinary knowledge production contexts.

To aid in the analysis and interpretation, Figure 5 presents all three layers alongside each other. It is constructed from the analysis of the lived experience and perceptions layer, and Table 22, Table 23, Table 24, and Table 25 that represent a summarized view of the technological and organizational layers of access tools and repositories. Figure 5 represents a generic performative model based on the understandings from the interviews with the researches and the documentary evidence. The links labeled as 1, 2, 3 and 4 in circles denote performative relationships between the three layers and researchers' individual knowledge production contexts. Link 1 denotes the performative relationship between the technological and the organizational layers, and it will be further described based on how the technological layers (see Table 22, Table 23) and the organizational layers (see Table 24, Table 25) inscribe their properties on each other.

Link 2 denotes the performative relationships between researchers' individual knowledge production contexts and the organizational layers. Link 3 denotes the performative relationships between researchers' individual knowledge production contexts and the technological layers. The cloud represents the disciplinary knowledge production contexts within which researchers' individual knowledge production contexts are enacted as they interact with the open access repositories, resulting in the lived experiences and perceptions layer denoted by Link 4. For example, the perception by the astronomers that arXiv and ADS are almost complete proxies to their disciplinary knowledge network is impacted by the inclusiveness (many sources and artifacts types) and integration (between ADS and arXiv, as well as data and astronomical object repositories) that was enabled at the organizational layer with the supported from the technological layer and various policies aimed at accelerating the rate of exchange of scholarly materials.

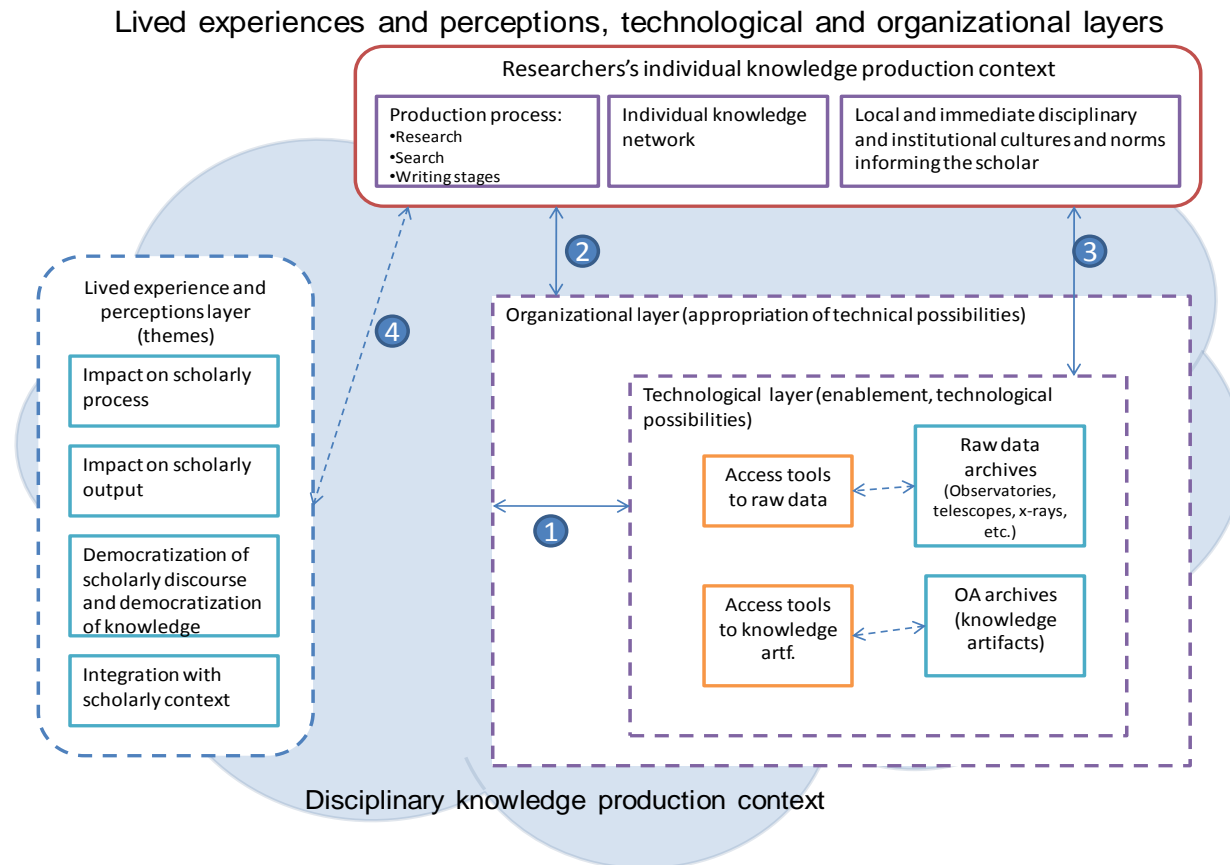
More specifically, Figure 5 is valuable in the analysis and interpretation of the contextual performative relationships between the three layers with respect to researchers' individual knowledge production contexts and the disciplinary knowledge production contexts. It also reveals that the local actors that co-construct each layer are also implicated in the co-construction of the localized individual knowledge production contexts. The co-construction dynamic has emerged as performing in two distinct ways: (1) realigning the existing relationships (social and technological) among the actors, and (2) the emergence of new relationships and actors (both local and global). Further, as it has been shown, in addition to the local actors (such as the researchers and their research processes and individual knowledge networks), global actors from the disciplinary

knowledge production contexts (such as the organizers and the managers of PhilSci, arXiv, and ADS, and tenure committees) are implicated in the co-constructed of the technological and organizational layers.

The four links in Figure 5, as well as the co-construction dynamics within the layers and between the layers, are described and interpreted in more detail in this chapter as they relate to Research Questions.

This chapter ends with the analysis and interpretation of the co-construction dynamics of the knowledge production contexts at individual level as well as and the co-construction dynamics of the disciplinary knowledge production contexts. The co-construction dynamics are related, compared and contrasted through the corresponding properties and dimensions of the lived experiences, technological and organizational layers for the two groups of researchers.

Figure 5: Performative model based on lived experiences and perceptions, technological and organizational layers



Notes: The technological and organizational layers are based on documentary evidence. The lived experience and perceptions layer is based on the interviews with the researchers. Links (1), (2) and (3) represent the performative relationships between the layers. Link (4) represents the lived experiences and perception layer.

### *12.1 Answering the Research Questions*

In this section, the analysis and findings from Chapter 8 through Chapter 11 are contextualized to address the Research Questions as set forth in section 3.1.

RQ1: How do researchers experience and perceive the role and value that OA repositories and access tool provide in their knowledge production process?

This research question, RQ1, has been addressed in Chapter 7 through Chapter 10. The lived experiences and perceptions of the astronomers and philosophers of science as they interact with the open access repositories and the access tools have been analyzed and interpreted at individual levels as well as at group level as they relate to the four themes that emerged iteratively from the open and axial coding of the interviews with the researchers. The analysis and interpretation of each scholar's lived experiences and perceptions have been summarized in tables that relate the perceived properties of the access tools and the repositories with respect to the four themes. The tables further relate the properties with the value they are perceived to provide and the role they are perceived to play in researchers' knowledge production processes and knowledge networks. The properties, values and roles with respect to the four themes have been aggregated and summarized for both groups of researchers (in sections 10.1 and 10.3 respectively) in order to understand and derive at lived experiences and perceptions at group level. The perceptions and lived experiences between the astronomers and the philosophers of science are compared and contrasted in section 10.5.

As it has been shown earlier, both the astronomers and the philosophers of science commonly perceive the value and the role of the open access repositories and the access tools to emerge from their properties as: a) enablers of openness, b) enablers of early,

quick and fast access to scholarly materials, and c) enablers of wider dissemination of scholarly material. However, beyond these common perceptions about the role of open access resources, there are distinct differences in how the arXiv scientists perceive the value of arXiv from the way the PhilSci scholars perceive the value of PhilSci as they are related to the four themes.

Even the property of openness, as the most ubiquitous property that has emerged from the interviews, is perceived to have different implications at the technological and the social levels for researchers' individual knowledge production context. With respect to the perception as a remover of barriers to entry both groups of researchers perceive that open access repositories enable new researchers to enter the disciplinary scholarly discourse (especially those from smaller institutions), and that open access repositories enable new research findings to enter the disciplinary knowledge network immediately (as pre-prints and manuscripts) without having to rely on the commercial journals for distribution. However, while the arXiv scientists articulate the enablement of interactions at the technological level (perceived to increase articles' visibility, discoverability and accessibility) as well as the social level (as enabler of collaborations between scientists from different institutions), PhilSci scholars articulate the value of the interactions as enabler only at the social level—enabling collaborations amongst scholars. While arXiv scientists turn inward by associating the technological level properties with the values provided by OA repositories and perceive the value of openness as performing on arXiv's capabilities to enhance the search, discovery and accessibility to the knowledge artifacts (by enhancing the technological features to help the scientists through linking the various resources automatically), PhilSci scholars turn

outward and articulate the interactions only at the social level, including the complementary nature of Home Pages, invisible colleges, Google Scholar and JSTOR in their individual knowledge production context that enable scholars from different institutions to collaborate with each other. This finding, that arXiv scientists turn inward towards a centralized location while PhilSci scholars turn outward towards a more dispersed set of resources is congruent with Carol Palmer's (1994) findings about the differences between the information practices of scientists and humanities scholars. Further, the emergence of ADS as a central and specialized access tool for astronomers, and the lack of such specialized and central access tool for philosophers of science is congruent with Jenny Fry's (2006) findings that disciplinary communities that exhibit high degree of mutual dependence in conjunction with low degree of research problem uncertainty (such as astronomy and astrophysics), are more likely to construct digital tools that will make their information practices more productive and more efficient. The lack of specialized and central access tool for philosophers of science (as participants in a discipline that exhibits low degree of mutual dependence and high degree of research problem uncertainty) is also congruent with Fry's findings that centralized and interconnected digital resources are not necessarily co-constructed in disciplines that exhibits low degree of mutual dependence and high degree of research problem uncertainty.

Derived from the analyses and interpretations so far, including my notes and memos, Table 26 shows the high-level comparison of the lived experiences and perceptions based on the interviews with both groups of researchers. The arXiv scientists perceive arXiv and ADS as highly performative on the scholarly context (i.e., they trust arXiv and



ADS to provide almost a complete view of the disciplinary knowledge network), with high impact on their scholarly knowledge production process (enabling more efficient search and research process) and on their individual knowledge networks (easy and efficient identification of knowledge artifacts for local use relevant for the production of a specific knowledge artifact), as well as being highly performative on the democratization of the scholarly discourse by removing the barriers to entry, and thus, enabling scholars from smaller institutions and unpublished or non-normative knowledge artifacts to enter the scholarly discourse. PhilSci scholars on the other side perceive very low integration with the scholarly context, limited to using PhilSci as a collaboratory for the exchange of ideas and a preparation tool for conferences, although they perceive that they can still do the same without PhilSci's help. PhilSci scholars perceive the PhilSci archive to perform mostly on their scholarly process, although at medium level, by enabling them to find and access immediately the latest research findings in their discipline. Even here, discovering and accessing the latest articles is not necessarily an exclusive process enabled by PhilSci. As it has been explained earlier, it is rather augmented by the Home Pages of other scholars, as well as their networks of invisible colleges.

Table 26: High level summary and comparison of lived experiences and perceptions

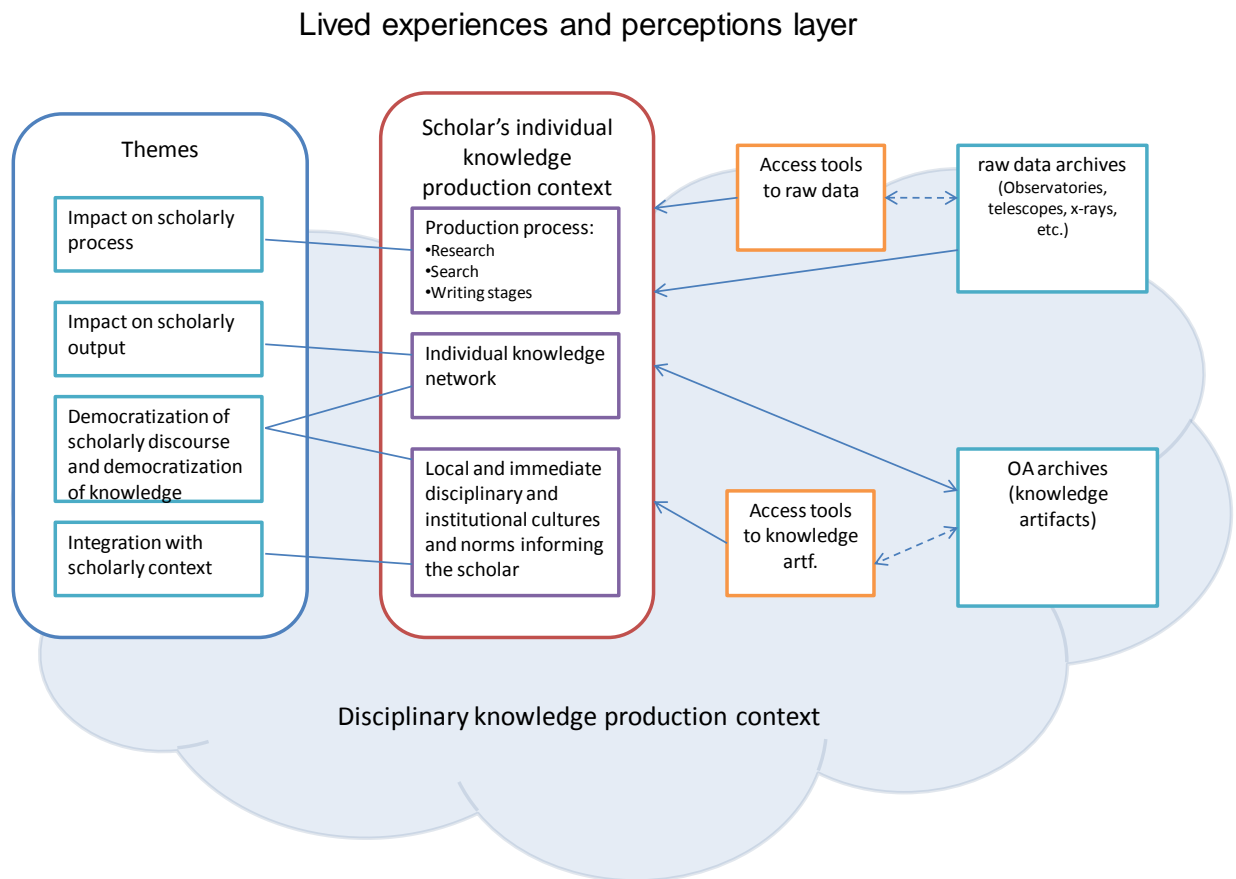
Themes	Dimensions as acknowledged by the researchers related to each theme		Knowledge production contexts
	arXiv scientists	PhilSci scholars	
Integration with scholarly context	High	Low	Disciplinary
Impact on scholarly output	High	Low	Individual
Impact on scholarly process	High	Medium	Individual
Democratization of scholarly discourse; barriers to entry removed	High	Low	Disciplinary

Source: Interviews with the researchers.

Table 26 also shows the mapping between the four themes of discourse with the level of knowledge production context as experienced and perceived by the researchers. The perceived impact on scholarly output and scholarly production process is locally and immediately related to researchers' individual knowledge production contexts. The integration with the broader context (i.e., the disciplinary knowledge production context) is enabled via the open access repositories and the access tools acting as links to enable the translation of disciplinary level knowledge production context for local use by the researchers. The performative capabilities are enacted by the researchers as they interact with the repositories and the access tools, to select a subset of the disciplinary knowledge network for localized use for the production of a specific knowledge artifact. In comparison to PhilSci, arXiv has higher performative ability (centrally situated and used) in translating the disciplinary knowledge network for local use by individual researchers. PhilSci's performative ability is low from the perspective that it can translate only a subset of the disciplinary knowledge network (from the corpus of articles that is available in PhilSci) for individual use by the philosophers of science. For the philosophers of science, in addition to the PhilSci archive, the translation of the disciplinary knowledge network into a subset for local use by individual scholars is augmented by the performative abilities of Home Pages of scholars, invisible colleges and JSTOR (perceived as delayed open access repository). The translation also occurs from the individual (local) to the disciplinary as scholars deposit their produced knowledge artifacts into the open access repositories that in turn can be accessed globally by other interested scholars.

A graphical representation of the relationships in Table 26 is shown in Figure 6. It shows the relationships between the four themes, researchers' individual knowledge production contexts, and the corresponding access tools and repositories.

Figure 6: A model for lived experiences and perception layer based on the interviews with researchers



The dynamics of researchers' interactions with the open access repositories are shown in Figure 6 as part of the larger disciplinary knowledge production contexts of their disciplinary fields, denoted by the cloud labeled "Disciplinary knowledge production contexts." Researchers' individual knowledge production contexts are linked to the respective disciplinary knowledge production context via their participation in the

respective disciplinary culture. Thus, the goal of Figure 6 is to show that researchers' individual knowledge production contexts are interrelated with the broader disciplinary context via the performative abilities of the repositories and the access tools that translate back and forth between the individual researchers (local) and the disciplinary level (global).

RQ2: How are the properties of the organizational and technological layers of OA repositories and access tools implicated in researchers' individual knowledge production contexts?

This research question, RQ2, has been addressed in Chapter 11 based on the analysis of the documentary evidence. The open access repositories and the corresponding access tools have been analyzed with respect to the features and functional capabilities that are provided by the software that is used to build them as technological systems. Emerging from the documentary analysis, following are the main features and functional capabilities that are provided and enabled by the software components that build arXiv, PhilSci, ADS and Google Scholar: a) support for multiple file formats, b) support linking between abstracts and full text, and traversing back and forth from articles to the references listed in the bibliography, c) provide searching and browsing capabilities, d) support meta-data for description of scholarly artifact and interoperability via protocols for meta-data exchange (OAI and OAI-PMH), e) support different types of artifacts in addition to pre-prints and post-prints (such as abstracts, technical reports, etc.), f) provide mechanism for daily updates via e-mail, g) provide mechanism to organize the submitted materials into meaningful categories, and h) enable harvesting of materials

from multiple sources. The features and functional capabilities between the technological systems that enable the structuring of PhilSci, arXiv and the access tool as organizational structures have been identified and compared in Table 22 (technological level for repositories) and Table 24 (technological level for access tools). The properties of the technological layers manifest themselves as features and functional capabilities that may be inscribed into the organizational layer to build arXiv and PhilSci as socio-technological information environments.

Similarly to the identification of properties at the technological layer, the intended roles and values of the open access repositories and the access tools (described in the documentary evidence) as articulated by their organizing and management structures, have been identified and compared in Table 23 (organizational level for repositories) and Table 25 (organizational level for access tools). The main values and roles of the repositories and the access tools as intended by the organizing and management structures for use by the scholarly community are as follows: a) enable open access to scholarly materials, b) become central location where researchers associated with the discipline can find scholarly materials, c) supplement the traditional scholarly process by enabling faster and immediate access to the latest research findings, d) enhance researchers' research process by making it more efficient, e) provide searching and browsing capabilities, f) establish a mechanism to ensure only relevant materials are deposited (no quality control through peer-review), g) in addition to other scholarly artifacts support conference submissions, h) provide permanent access to the deposited materials, i) enable access to multiple sources, j) provide complete access to the

knowledge artifacts that define the discipline, and k) be a valuable resource for tenure committees.

The findings reveal that the technological layer properties of the open access repositories are almost the same for arXiv and PhilSci, with one difference that arXiv's automatic linking capability between citations and bibliographies across the articles deposited therein is extensive and sophisticated, and used greatly by astronomers and astrophysicists on a daily basis. The distinction between the technological and the organizational layers of the access tools substantially differ between ADS (access tool for arXiv) and Google Scholar (as a generic access tool for PhilSci). ADS extends the linking capabilities of arXiv by introducing abstracts, raw data, data catalogs, and indices into the corpus of materials that can be automatically linked. Thus, the technological layer properties have been inscribed and performed onto the organizational layer. For the astronomers the interaction at the technological layer (linking between resources) is one of the main dimensions associated with the property of openness. It enables them to easily traverse between the corpuses of materials and quickly identify the knowledge artifacts they need for the research problem at hand. The version of EPrints that powers PhilSci does not provide such capability, thus, the linking between resources is left to be attempted manually by the philosophers of science.

As it has been described in more detail in sections Chapter 11, arXiv and ADS have been co-constructed over time to include almost all of the research output in the discipline. Thus, arXiv and ADS have emerged as trusted location independent central actors that are perceived as a proxy to the complete body of knowledge artifacts that define astronomy and astrophysics. This has in turn performed on the information work

of astronomers by substituting arXiv and ADS for the library and journal portals. It has also enabled astronomers to view easily and efficiently a subset of the disciplinary knowledge network, by building a localized knowledge network (through searching and downloading) that is needed for the production of knowledge artifacts. The property of inclusiveness was not very strongly perceived by the philosophers of science—their view of open access is more distributed in nature with resources scattered across PhilSci, Home Pages, JSTOR and the invisible colleges, as a consequence of which philosophers of science spend more time performing manual tasks across different resources and interfaces.

Further, the properties of the organizational layers of the open access repositories and the access tools are perceived by the researchers to have performed on their individual knowledge production contexts by “shrinking” time and space in their production process, thus enabling them to use their time more efficiently and effectively for research instead of searching for knowledge artifacts. Further, the “shrinking” time and space has performed on researchers’ construction of their individual and localized knowledge network by bringing together knowledge artifacts from different sources and different publication periods in close proximity to each other. The researchers thus perceive that the implication of the open access repositories and access tool on their production process and their individual knowledge network has performed on the type of research problems they address and the subset of knowledge artifacts they consider, that in turn impact the quality, comprehensiveness and scope of their scholarly output.

RQ3: How are the actors and properties of the broader socio-cultural and technological context, including disciplinary norms and cultures, implicated in researchers' knowledge production context?

In addition to the properties of the access tools and the repositories, the researchers also perceived the broader context to be implicated in their scholarly knowledge production context. The arXiv scientists perceive that the trust in arXiv and ADS with a central role in astronomy and astrophysics has performed on their search process by almost removing library and journal portals from the information practices and replacing them with ADS and arXiv. Considering that over time PhilSci does not seem to have emerged with a central and an independent role in the philosophy of science discipline, PhilSci scholars perceive that it has a minor role in their knowledge production processes, with strong perception that PhilSci is most important for finding the latest research findings ahead of time (before publication in the commercial journals).

For the arXiv scientists the translation (i.e., aggregation of resources in central location) capabilities are inscribed into the organizational and technological layers of the open access repositories and access tools, while for the PhilSci scholars the aggregating of resources is enacted by the scholars themselves by visiting PhilSci, scholars' Home Pages, utilizing invisible colleges, and Google search—where the PhilSci archive acts as only a partial aggregator to enact the individual knowledge networks of PhilSci scholars. The perceived property of trust has thus been co-constructed over period of time. At the moment (arXiv has been in operation since 1991 and ADS since 1988), the trust in arXiv and ADS to provide the latest and almost complete access to disciplinary knowledge in astronomy and astrophysics is perceived to be very high. Thus, the arXiv scientists



perceive the combination of arXiv and ADS as central in how they enact their individual knowledge networks, visiting for the most part only ADS in their search process since ADS has access to all of the records in arXiv as well as other resources such as abstracts, technical reports, links to raw data and astronomical objects, etc. PhilSci scholars however, due to PhilSci's low perception of trust value, perceive PhilSci as complementing their existing means of access to knowledge artifacts. For the PhilSci scholars the open access resources are dispersed across different contexts, where Home Pages of scholars act as a distributed open access repository (sometimes linking to each other with manual links) and JSTOR is perceived as a delayed open access repository that institutions can subscribe to at low rates (JSTOR subscriptions do not include the last 3-5 years of publications).

In addition to arXiv and ADS, arXiv scientists need access to raw data for their knowledge production. The raw data are produced by observatories, telescope and other instruments that gather and record data related to astronomical events. Thus, in conjunction with access to arXiv and ADS, astronomers and astrophysicists need to access the raw data repositories and various data catalogs on daily basis and they are as critical for their knowledge production as arXiv and ADS. The tenure process as an institutional and disciplinary actor is also perceived by arXiv scientists to be relevant in their interaction with arXiv. They perceive that tenure committees heavily rely on ADS as a tool to assess scientists' impact and thus make sure to have their pre-prints and post-prints be visible, discoverable and accessible through ADS.

The length of time that the repositories have been in operation does not completely explain the perceived trust by the researchers. The interviews with the arXiv scientists

reveal that inclusiveness (how well do these repositories represent the disciplinary knowledge ecosystem) is the central property that has performed on researchers' perception of trust. The emergence of ADS and arXiv as trusted resources is strongly associated with the cultural configuration of a discipline and researchers' information practices, as explained through the frameworks of Bates (1994), Fry (2006), Knorr Cetina (1999) and Palmer (2005), that inform the use and the nature about how information technology artifacts are appropriated and adapted both at individual and disciplinary levels. Thus, the open access phenomenon, instantiated in scholars' research process as open access repositories and access tools, is only one of the actors that is implicated in the socio-technological co-construction dynamics of the knowledge production context by enhancing articles' visibility, discoverability and accessibility.

RQ4: How are the performative agencies of the key actors and contexts implicated in researchers' individual knowledge production context?

The performative agencies of the different actors that emerged from the interviews as well as the documentary evidence have been described and elaborated in Chapter 8 through Chapter 11. Here, the performative agencies of the central actors as perceived by the researchers to have impacted their knowledge production process and knowledge networks are addressed. As it has been already mentioned, the property of openness has emerged as a comprehensive property more directly defined through its performative abilities perceived to remove barriers to entry and enable new interactions. With respect to the ability as an enabler of new interactions, both arXiv and ADS have been

constructed with capabilities and features to automatically link the resources via their citations and bibliographies further increasing articles' discoverability.

In addition to the key property of openness as the main property that has performed on the knowledge artifacts increasing their visibility, discoverability and accessibility, other key actors and properties that have been perceived by the researchers to have performed on researchers' knowledge production process are: a) the efficiencies by which the open access repositories and the access tools have enabled researchers to discover and access scholarly materials (browsing, searching, traversing through bibliographic links, strong integration), b) the ability to access materials from different sources in one location, c) the ability to access the latest research findings immediately and before they are published in the commercial journals, and d) enabling researchers to connect to each other where open access repositories are perceived as collaboratories for work in progress.

The availability of materials from many different sources in one location has performed on scholar's search processes by making them more efficient in finding the materials needed for their *knowledge production process*. As a result, arXiv scientists stated that for the most part they no longer visit commercial journals' portals or their library portals. This is an example of how researchers' research processes were restructured and realigned to substitute a number of different resources with one that is centrally managed. The openness of the meta-data and the open protocols and open standards, have enabled citation and bibliographic linking as well as linking between arXiv and ADS and between abstracts and the full texts of pre-prints or post-prints. The linking capability has performed on the scholarly production process by realigning

researchers' research process with the introduction of process steps that researchers can undertake themselves. Thus, based on researchers' perceptions two types of realignments have emerged: a) that their traditional research process was enhanced by modification, where the same process steps became more efficient, and b) by the introduction of new approach to the research process made possible in the digital realm—ability to scope the research problem with ease by using the linking capabilities provided by the open repository and the access tools.

The main contextual actor that has been perceived by the researchers to have had performed on their knowledge work and as well as their knowledge networks, has emerged from the disciplinary culture as a perceived trust by the researchers in whether the repositories and the access tools are relevant and reputable in their discipline. Trust has already been discussed in RQ3.

Considering that researchers and the tools at their disposal are part of the broader disciplinary knowledge production context, the dynamics of the co-construction are instantiated and informed by the knowledge production context. By the dynamics of co-construction it is meant that there are three distinct layers that are each internally co-constructed by performing on each others' properties by inscription; and that the features and technological capabilities available at the technological layers are inscribed by appropriation into the organizational layers (link 1 in Figure 5, such as the use of linking capabilities to traverse from one article to another and into the raw data and astronomical objects), and that the intended organizational values and roles and the available technological features and functional capabilities perform on researchers' perceptions and lived experiences (link 2 and 3 in Figure 5 respectively, such as the

perception that the linking capabilities have enhanced articles' discoverability). This represents the performative feedback loop that is instantiated over time.

RQ5: How are the organizational and technological layer properties of the open access repositories and the access tools related and associated?

This research question, RQ5, has been addressed in Chapter 11 based on documentary evidence. The technological and organizational properties of the access tools and the repositories have been analyzed and interpreted in sections 11.1 and 11.2. The relationships and the dynamics between the technological and the organizational layers are analyzed and interpreted in section 11.3. To summarize the relationships between the layers and their properties, it is important to note the different nature of the properties in the three different layers. In the technological layers, the properties are described as features and functional capabilities, shown in Table 22 and Table 24, made available by the software that is used to build the technological layer of the open access repositories and the access tool. The OAI-PMH protocol and standards for the open exchange of meta-data have performed on the structuring of the technological layers by inscribing their features (i.e., being incorporated) in the construction of the software. In the organizational layers, the properties are described as roles and values, shown in Table 23 and Table 25, intended to help the scholarly community and especially to help the researchers that use the open access repositories and the access tools in their knowledge production processes. Comparing repositories' technological features and capabilities (see Table 22) with repositories' organizational intended roles and values (see Table 23) of both arXiv and PhilSci, reveals that the technological layer performed

on the organizational layer by inscribing the technological features and functional capabilities in conjunction with the inscription of policies that are also inscribed into the organizational layer, such as the policies that define what materials are included and with which systems should be integrated. For example, the technological capabilities of arXiv and ADS for searching, browsing, support for different types of file formats and different types of artifacts, have been made available at organizational layer for the researchers to use them. Thus, at the organizational layer, arXiv and PhilSci are at least performed by properties from the technological layer as well as the policies enacted by the administrators and management of arXiv and PhilSci respectively that determine what technical capabilities are made available for use, what materials can be included and how the repositories will be linked to other systems (such as ADS). The distinction between the technological and the organizational layer is clearer in PhilSci than arXiv. PhilSci administrator and management structure may not be able to enact a policy of interoperability with other systems and inclusiveness of materials from different sources unless it can be supported by EPrints. To give another example, the organizational intent for both arXiv and PhilSci to be able to support different types of artifacts (pre-prints, post-prints, conference submission, etc.) has been enacted either explicitly or implicitly by the organizing and management structures of the respective repositories. However, this intent had to be supportable by the technological layers. Thus, there is indication that the relationship between the organizational and the technological layers is co-constructionist in nature, situated in and performed by the disciplinary context, where the layers perform on each other by inscribing and modifying each others' properties in

order to provide users with capabilities to enhance their knowledge production processes (at individual and disciplinary level).

The same co-constructionist dynamic that has emerged between the technological and the organizational layers of the open access repositories, has also emerged between the technological layers (see Table 24) and the organizational layers (see Table 25) of the access tools—the technological capabilities are instantiated into organizational capabilities as intended by the organizing structures and management of the access tools. For example, the ability for astronomers to search multitude of sources has to be supported by the technological layer. Because of the bibliographic linking capabilities (that enhances articles' visibility, discoverability, and accessibility) that have been made possible via the tight integration between arXiv and ADS, and because of the almost complete availability of all of the resources relevant to astronomers and astrophysicists via arXiv and ADS, scientists perceive that ADS has become one of the main tools that tenure committees use to assess scientists' citation impact. Here we see how the emergence of an organizational layer property (i.e., use of ADS for citation impact assessment), that is not clearly expressed in the documentary evidence as a value that ADS provides, has been perceived by the arXiv scientists as being appropriated by tenure committees as being used in the tenure assessment process.

To further clarify and contextualize the scope of the co-construction nature between the technological and organizational layers, Figure 7 presents the relationships between the technological and organizational layers that are implicated in the production processes of both groups of researchers and their mutual performative processes with their respective disciplinary contexts. The distinction between the construction of the

technological and the organizational layers of arXiv and ADS, and PhilSci and Google Scholar is that arXiv and ADS are structured within the disciplinary context, while only the organizational layer of PhilSci is structured within the disciplinary context, with EPrints and Google Scholar being constructed outside of the disciplinary context. This distinction is congruent with Fry's (2006) findings that in this specific case would categorize the use of EPrints and Google Scholar as external for the philosophers of science:

The lack of centralised coordination and control in these fields [such as philosophy of science] will make it difficult for the scholarly community to systematically appropriate and develop digital infrastructures and resources in response to specific cultural needs. Often such fields have to work within externally imposed and developed digital infrastructures and resources (p. 312)

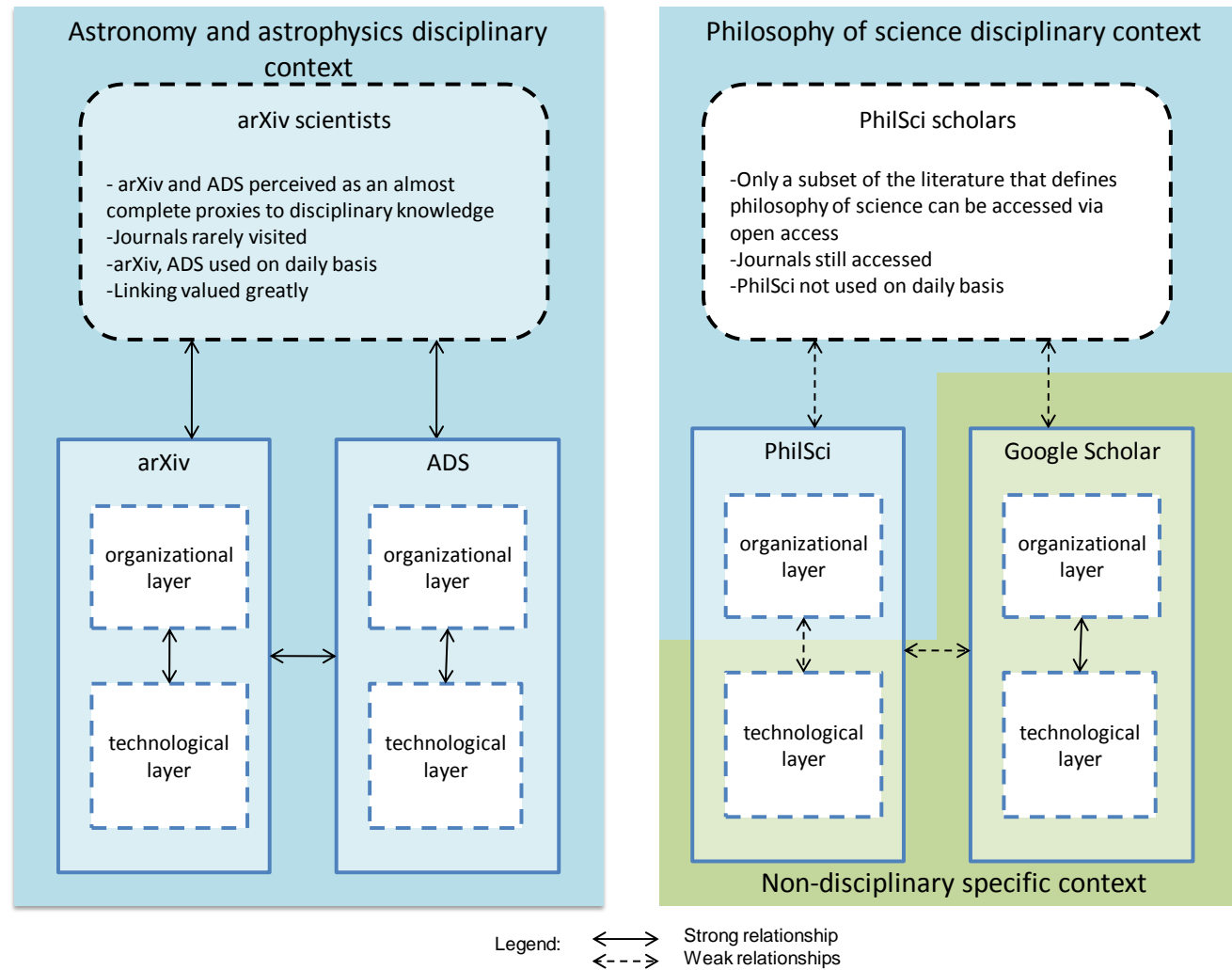
Therefore, arXiv's and ADS's strong performative capability on astronomers' knowledge production process, and the weak performative capabilities of PhilSci for the philosophers of science, can be partially explained by the property of "openness" in the co-construction of the technological layer of open access resources. The linking, relating and associating of resources for arXiv is achieved at the technological layer, suggesting that disciplines that exhibit high mutual dependence and low problem uncertainty (such as astronomy and astrophysics) are more congruent with offloading and assigning procedural and systematic task to be automated by digital resources and tools, with open access (as a remover of barrier to integration and inclusion) enhancing and strengthening the co-construction of arXiv and ADS. On the other side, a disciplinary context that exhibits low mutual dependence and high problem uncertainty is not as congruent with



assigning its tasks for automatic and systematic processing by digital tools because they are too complex to be attempted, and the property of “openness” cannot significantly affect the outcome of appropriation and construction, such as the case with PhilSci.

To further contextualize within Fry’s framework, the property of openness is a relevant actor in the co-construction of open access resources in a specific discipline, but the nature of its implication is mediated by disciplinary characteristics such as those put forward by Whitley (2000) upon which Fry has build by extending Whitley’s framework in the digital realm. To use the language of ANT, open source resources become one of the translating actors through which the role and the implication of a disciplinary culture on researchers’ individual knowledge production context can be understood and explicated.

Figure 7: Relationship between disciplinary culture contexts and the technological and organization layers



RQ6: How are the open access repositories and access tools implicated in the structuring of the knowledge production contexts of the arXiv scientists and the PhilSci scholars?

To understand the difference between the interaction dynamics that emerge as arXiv scientists and PhilSci scholars interact with the respective open access repositories and access tools, the interactions between the technological and the organizational layers and between the organizational layers and the lived experiences and perceptions layers are described with respect to the properties and across their dimensions.

As it has been shown previously by referring to Table 22, Table 23, Table 24, and Table 25, the interactions between the organizational layers and the technological layers exhibit co-constructionist dynamics. The organizational layers are a visible part of the broader context that defines scholarly communication and more specifically they are a visible part of the disciplinary cultures depending on their level of integration in the scholarly process. The technological layers are part of the broader context of software development and more specifically the development of open source software that is available at no cost. The properties that define the actors from the technological layers and the properties that define the actors in the organizational layers mutually perform on each other resulting in the construction of open access repositories and access tool as socio-technological structures that are used in researchers' knowledge production process.

Even though the arXiv scientists differ from the PhilSci scholars with respect to what access tools have been made available for their use (see Table 24 and Table 25), the organizational structures that have build arXiv and PhilSci describe their intended goals

by articulating and describing almost the same values (see Table 22 and Table 23) that these tools can provide to the scholarly community, such as browsing, searching, support for multiple file format and artifact types, open access, supplement the traditional scholarly process, enhance researchers search process, etc. Thus, we have two sets of researchers, astronomers and philosophers of science, using the repositories made available in their discipline, and yet, based on their responses to the interview questions they greatly differ in their perception about the role and value the open access repositories play in their individual information practices as well as the role they play within the discipline. These differences are elaborated next.

The interaction dynamics between the technological and organizational layers of the open access repositories and the access tools differ between the open access resources used by the arXiv scientists and those used by the PhilSci scholars—as has been explained in RQ5, the technological layer of PhilSci (i.e., EPrints) is constructed outside of the discipline and therefore has weak performative power on the organizational layer of PhilSci. The arXiv archive has emerged as an actor that translates the dispersed disciplinary knowledge ecosystem (pre-prints and post-prints from multitude of journals) into an inclusive, aggregated, and a well-managed structure where the knowledge artifacts are automatically linked and associated through their citations and references. ADS, as an access tool, has further strengthened the inclusiveness by aggregating data catalogs, raw data repositories, indices, and abstracting services, and integrating and linking them with the knowledge artifacts available in arXiv. For the PhilSci scholars, the PhilSci archive has emerged as only one of the actors that enable scholars to search, discover and access open access articles. Searching for the latest articles is usually done

either by using directly the websites of PhilSci, JSTOR and the Home Pages of scholars, or by using generic (for topical perspective) search engine such as Google Scholar.

As it has been shown earlier, the arXiv scientists perceive the implications of open access resources as removers of barriers to entry, enabling scientists to enter the scholarly discourse by making the knowledge artifacts widely available and accessible free of charge, and also enabling new knowledge to enter the disciplinary knowledge ecosystem much earlier. This has been perceived to have performed on scientists' collaboration practices by enabling them to discover other scientists with similar interests in order to collaborate together. Further, the arXiv scientists turn inward in their articulation of arXiv and ADS and perceive that arXiv's and ADS's central role (enabled at technological level by automatic linking, open integration and the broad inclusiveness of scholarly materials) has realigned their research process as they have replaced their traditional sources of knowledge artifacts such as library and journal portals. PhilSci scholars on the other side articulated the implications of open access mostly as a remover of barrier to entry that has performed on their research process, enabling them to read the latest research findings immediately by reducing the publication gap, enabling collaboration with other scholars, where PhilSci is perceived as a collaboratory for work in progress and as a preparatory and publication place for conference papers. While arXiv scientists perceive the open access resources with a central role in their discipline, PhilSci scholars perceive PhilSci as one of the actors in their production process. Home Pages of scholars, JSTOR, and the invisible colleges have performed on the PhilSci scholars' knowledge production processes in a distributed fashion by augmenting their search processes.

The performative capabilities of arXiv, ADS and PhilSci on the respective groups of researchers are presented graphically in Figure 8 and Figure 9, as models of researchers' perceptions of the context in which the digital tools are integrated with their work practice.

As shown in the socio-technological model of astronomers' knowledge work impacted by open access (Figure 8), arXiv and ADS have a central role in the knowledge production context of the arXiv scientists. For the most part, they perceive that the disciplinary knowledge network of astronomy and astrophysics is almost fully visible, discoverable and accessible via ADS and arXiv. Thus, arXiv and ADS have realigned scientists' knowledge production process by almost removing journal and library portals from their research process. These open access resources, as intermediaries between the scientists and the knowledge artifacts, perform on the disciplinary knowledge network as a filter through which scientists are able to create a localized knowledge network to be used in the production of articles. In comparison, as shown in the socio-technological model of philosophers' of science knowledge work, PhilSci scholars perceive that only a subset of the disciplinary knowledge network is visible, discoverable and accessible via PhilSci. Thus, they complement searching process with JSTOR, Home Pages of scholars and invisible colleges. As already noted, PhilSci scholars value the PhilSci archive for its ability to enable immediate and fast access to the latest research in a pre-print form and its ability to increase the visibility of articles deposited therein. Thus, although PhilSci scholars use PhilSci in their scholarly process, they limit its value to discovering the latest research findings quicker and much earlier before they appear in the commercial journals.

Figure 8: Socio-technological model of the arXiv scientists

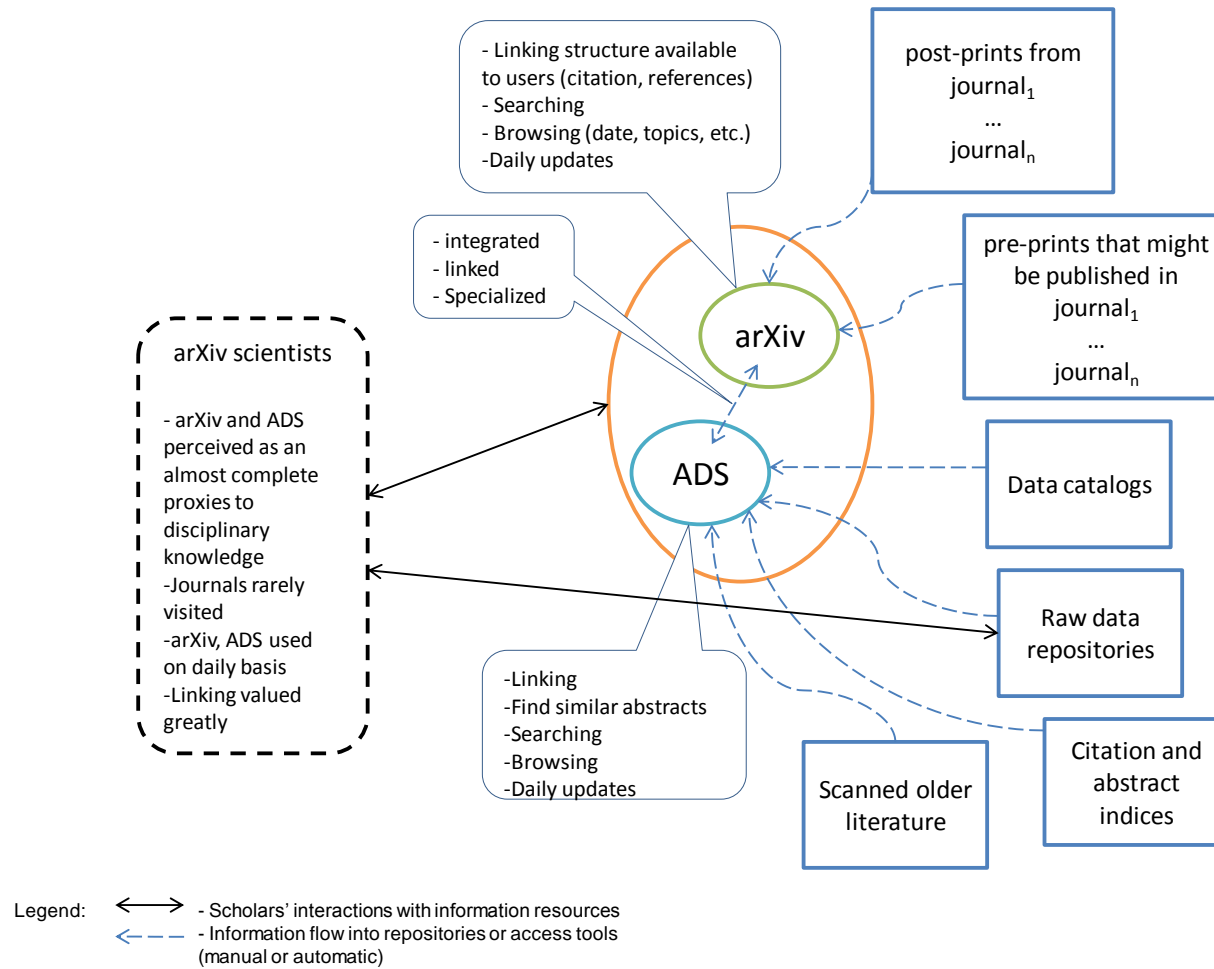
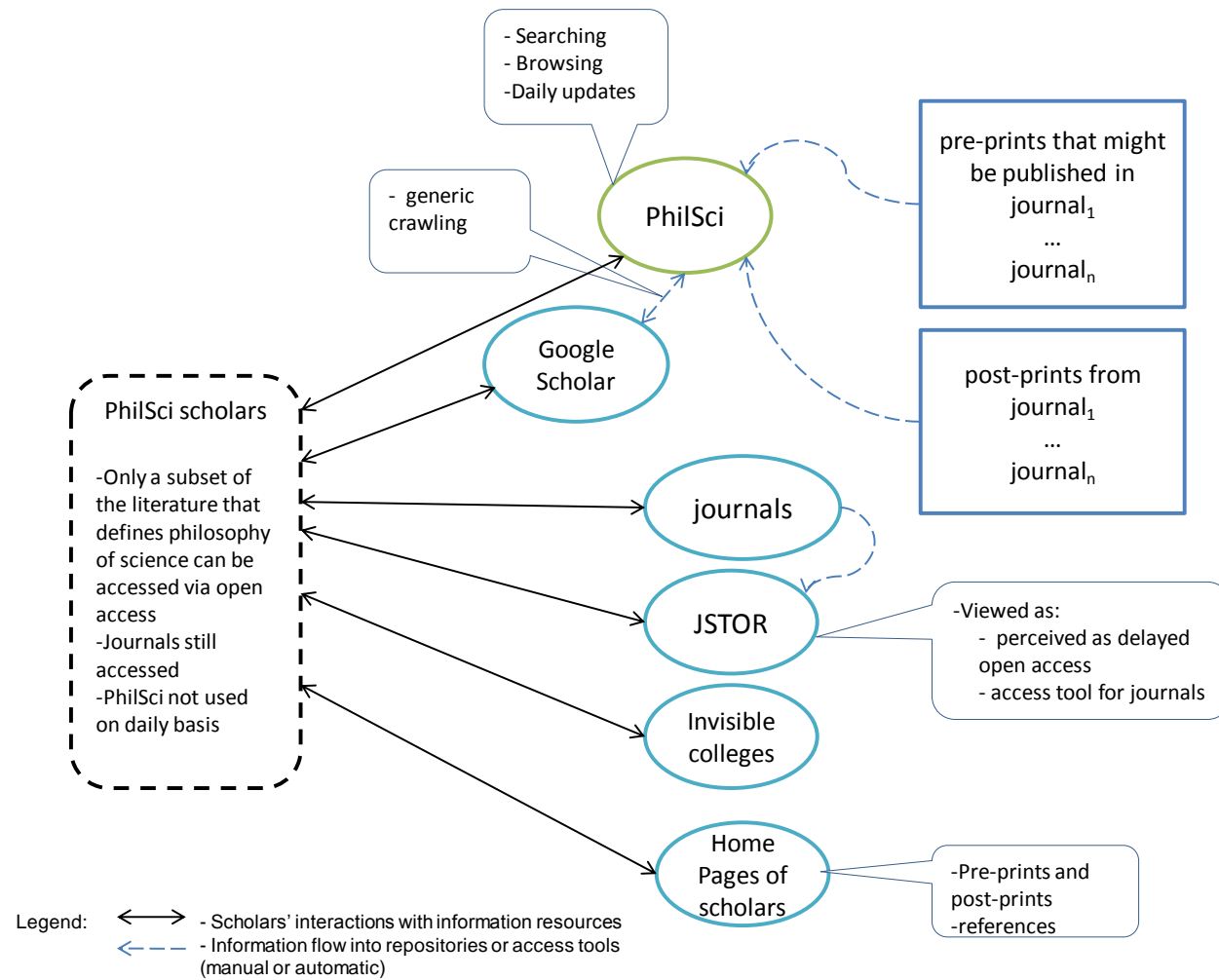


Figure 9: Socio-technological model of the PhilSci scholars





To summarize, as it can be observed in Figure 8, ADS and arXiv are the actors that centrally coordinate, manage and perform on various aspect of astronomers' research processes across multitude of resources that are integrated into arXiv and ADS. This finding is congruent with Palmer's (2005) observation that scientists' research processes are less uncertain and more defined: "Searching, collecting, and consultation are more targeted and endpoints tend to be more defined" (p. 1146), which is different from the work of humanities scholars whose research process is more "open ended" and requires more personal and closer reading and investigation of the literature.

The processes as presented in Figure 8 and Figure 9 are also congruent with Fry's (2006) theoretical framework relating the emergence of digital resources and communication channels (such as ADS and arXiv) in disciplinary fields that exhibit high level of mutual dependence and low level of task uncertainty.

### *12.2 Socio-technological implication for the four themes*

In this section, the findings are summarized and contextualized with respect to the four themes and their implication for the co-construction of researchers' individual knowledge production contexts and the corresponding disciplinary knowledge production contexts for astrophysicists and the philosophers of science.

The four themes emerged from the interviews with the researchers based on their perceptions about their interaction with the open access repositories and access tools. Considering that the access tools and the open access repositories have performed on researchers' information practices by realigning their knowledge production processes and knowledge networks, the knowledge artifacts that are produced as a result of these realignments make it back into the disciplinary knowledge ecosystem to be accessed by

other researchers. This cycle, a performative feedback loop, emerges as a primary mechanism for circulating scholarly knowledge utilizing the open access resources.

#### 12.2.1 Impact on knowledge production process

Open access has impacted the knowledge production process by realigning the existing processes and making them more effective and more efficient, and by the emergence of new tools directly related to the availability of open access resources.

For the arXiv scientists, arXiv, ADS and the raw data repositories are integrated in their everyday scholarly life. They note very strongly that the open access repositories and access tools have performed on their knowledge production processes by realigning the research steps, especially their approach to searching, discovering and accessing articles, by making them more efficient and more effective. For example, their search process has been streamlined by searching arXiv and via the specialized search engine ADS through a normalized interface, instead of having to visit multitude of websites or generic library portals, giving them more time to deal with the research problem. Their knowledge production process has also been impacted by enabling them to use tools that were co-constructed as a result of the emergence of open access. For example, the automated bibliographic linking capability, enabled by arXiv and ADS that are perceived by the astronomers as an almost complete proxy to the disciplinary knowledge ecosystem, has performed on astronomers research process to enable easy traversing through a number of articles based on subject of interest to find and access the relevant articles.

While astronomers stated that they are not be able to do their work without arXiv, ADS and the raw data repositories, the philosophers of science did not perceive the

PhilSci archive as imperative to their scholarship. They perceive the PhilSci archive with limited impact on their knowledge production processes, limited only to enabling the scholars to quickly find the latest research findings, much earlier before they are published in the commercial journals. PhilSci scholars augment their search process by visiting JSTOR and Home Pages of other scholars, and also rely on their invisible colleges. Therefore, while arXiv and ADS perform strongly on astronomers' knowledge production process as a centrally managed resource, philosophers' of science knowledge production processes is performed by different set of resources.

#### 12.2.2 Impact on knowledge output

In conjunction with the implications for their knowledge production processes, arXiv scientists perceive that the availability of arXiv and ADS has performed on their knowledge output by the integration and inclusion of multitude of scholarly sources and types of knowledge artifacts. They perceive that the articles they produce are of better quality, more complete and more comprehensive. In addition, they perceive that because of the availability of articles from many sources in one location, they have a better view of the research problem and often serendipitously discover patterns between articles due to their temporal and spatial proximity, coupled with the efficient and effective way of accessing them enabled by the technological layers.

PhilSci scholars however do not perceive that the PhilSci archive has had an impact on their knowledge output. However, this perception is subject for further study as PhilSci scholars perceive that the PhilSci archive has enabled them to access the latest research findings immediately and much earlier than it would have been possible by waiting for them to be published in the journals.

### 12.2.3 Integration with the broader context

The integration of the open access resources with the scholarly context is perceived very strongly by the arXiv scientists in comparison to the PhilSci scholars. For the astronomers and the astrophysicist, arXiv and ADS are centrally positioned, integrated and trusted as a proxy to the disciplinary knowledge. Not that they are trusted only by the scientists, tenure committees also rely on ADS for citation impact assessment. Needless to say, such integration takes time and many different actors have been brought together whose relationships have been reconfigured across the technological and organizational layers. As it has been shown earlier, based on documentary evidence, the organizing structures of the open access repositories and the access tools built these open access resources with the intention to help the advancement of scholarship in their respective disciplines.

The findings from documentary evidence show that as organizational structures both arXiv and PhilSci have been constructed with very similar (almost equivalent) goals in mind. They both have stated goals to enable researchers to search, discover and access scholarly materials of various types (but mostly pre-prints and post-prints) free of any cost, with the hope that the open access availability will enhance the rate of scholarly output and research findings.

However, the outcome differs greatly within the two disciplines. For the astronomers, arXiv and ADS are integral part of their individual knowledge production contexts as well as integral actors in the disciplinary knowledge production context. For the philosophers of science, PhilSci is somewhat integrated with their individual knowledge production contexts, but the integration with the disciplinary knowledge

production context is very weak. Both arXiv scientists and PhilSci scholars perceive that the traditional paper based pre-print distribution culture has performed on the structuring of their respective open access repositories by inscribing the property of “openness” into the digital realm. This seems to indicate that perhaps some other properties that are intrinsic to the specific disciplinary culture may be implicated in the appropriation process of open access repositories by a specific discipline. The interviews with the researchers and the documentary evidence do not provide sufficient data to discern the dynamics that can account for these differences. However, as it has been mentioned earlier, Fry (2006) has shown that the production and use of digital resources and the control of communication channels—both features are instantiated within the open access repositories as information environments that enable collaboration and communication amongst researchers—are dependent on disciplinary socio-cultural and epistemic properties. Thus, the finding in this study that open access repositories and access tools exhibit strong co-construction and appropriation dynamic within the community of astronomers, and very weak co-construction and appropriation dynamic with the community of philosopher of science, is congruent with Fry’s (2006) typology that astronomers exhibit high degree of mutual dependence and low degree of task and research problem uncertainty.

#### 12.2.4 Democratization of the scholarly communication

By and large, the astronomers and astrophysicists perceive that the availability of the open access repositories has brought structural change to the discipline by democratizing scholarly communication. Open access has thus performed on the scholarly context by enabling researchers and institutions to participate in the scholarly discourse even

though they might not be able to afford access to the respective commercial journals. None of the arXiv participants themselves is affiliated with institutions that cannot afford the main journals for their discipline, at least they perceive so. Themselves they mostly see the value of open access in the ability to enhance their knowledge production process (efficient and effective access to all knowledge artifacts, including pre-prints) and the enhanced article visibility, discoverability and accessibility. However, except for one of the researchers in this study, they made a point to stress the value of open access as democratizing the scholarly discourse.

In addition to the democratization of the process, there is a perception that by enabling open access and wide distribution for documents that do not get published in the commercial journals (such as pre-prints, conference papers, proposals, technical reports, etc.), the knowledge ecosystem is also being democratized by enabling the inclusion of knowledge artifacts into scholars' localized knowledge network that otherwise are not published in commercial journals. One scholar mentioned the value of the knowledge that is embedded in pre-prints that do not make it into commercial journals, not due to the quality of the scholarly work, but because of administrative issues such as number of pages or number of articles per issues, the close call and the personal nature of the peer-review process, timing, etc. Another scholar mentioned how some very important pre-prints in arXiv received a large number of citations even though they did not pass the peer-review process. Independently of the reason why they did not pass the peer-review process, or whether they were submitted for publication in a commercial journal or not, there is a perception that arXiv as an open access repository

has the potential to realign the disciplinary knowledge ecosystem by enabling unpublished knowledge to enter the scholarly context.

For the PhilSci scholars, the perception that open access resources have democratized the scholarly discourse by enabling scholars from smaller institutions to participate was mentioned only by one scholar. He was affiliated with a smaller school that could not afford to subscribe to the main journals relevant to philosophy of science.

### *12.3 Co-construction of the individual and disciplinary contexts*

Intertwined with these four themes emerged an analytical model of a scholar defined with respect to scholarly work that is situated within disciplinary norms and cultures. The disciplinary knowledge production context emerges as a construct comprised of: a) the information practices enacted by researchers of the discipline during their knowledge production process, b) disciplinary knowledge ecosystem, and c) disciplinary norms and cultural behaviors related to the discipline, as Fry (2006) has shown by applying Whitley's (2000) theory on the degree of mutual dependence and the degree of task uncertainty in the digital realm. At individual level, researchers enact immediate and locally situated processes and use only a subset of the disciplinary knowledge ecosystem. Thus, the individual knowledge production context emerges as a construct comprised of: a) the information practices enacted during scholar's knowledge production, b) scholar's individual knowledge network relevant to the discipline, temporary and local for the specific research problem, and c) researcher's individual norms and cultural behaviors related to the discipline where researchers might build and utilize digital resources locally.

From this perspective, the open access repositories and the access tools manifest themselves as socio-technological actors with implications on researchers' individual knowledge production contexts (by reconfiguring researchers' local production processes and local knowledge networks), as well as with implications on the respective disciplinary knowledge production contexts, as researchers integrate their locally enacted digital resources and processes into the disciplinary (i.e., global) practices by collaborating with other researchers, albeit with varying degree of performative ability between the arXiv scientists and PhilSci scholars. At individual level, the open access tools and repositories are being used by the individual researchers whose scholarly endeavor is part of the scholarship that is defined by the discipline. Thus, the open access tools and repositories are co-constructed with actors not only from the technological and organizational layers, but also with actors and properties that define disciplinary norms, practices and cultures, as has been shown by Whitley (2000) and Fry (2006) by looking both at socio-cultural (degree of mutual dependence) and epistemic (degree of task and research problem uncertainty) dimensions of a scholarly discipline.

Researchers' knowledge production processes and their knowledge networks are distinct actors that define researchers' knowledge production contexts. However, as an intertwined set of actors that perform on each other, they are mutually performed over time by the disciplinary norms and practices to which they also contribute. Researchers for the most part operate within the disciplinary norms and practices. The localized and individual knowledge networks are also co-constructed with the disciplinary knowledge ecosystem as researchers work on research problems and contribute the knowledge output in the form of scholarly artifact back into the disciplinary knowledge network.



To summarize, the findings suggest that researchers' individual knowledge production contexts are co-constructed with the disciplinary knowledge production contexts by mutually performing on each other. The co-construction dynamics are performed by actors from the technological and organizational layers as well as by the intentions of the organizing structures that built the open access repositories. Open access resources such as the open access repositories and the access tools have entered the co-construction dynamics because they are able to realign, enhance and accelerate scholarly knowledge exchange by increasing article visibility, discoverability and accessibility.

### **Chapter 13.      The value of ANT for this study**

As suggested by the actor-network theory (ANT), neither the researchers with their information practices, nor the open access repositories and the related access tools operate in isolation from their surroundings. As it has been explained and interpreted in the previous chapters, researchers' individual knowledge production contexts operate within the context of the scholarly exchange, suspended in the disciplinary knowledge production contexts alongside the access tools, the resources needed for their knowledge production, and the institutional and disciplinary norms and practices.

In Chapter 4 ANT was suggested as a congruent theoretical and methodological approach to study problems that exhibit socio-technological co-constructionist properties. Here the specific benefits of the ANT approach are described with respect to the theoretical and methodological implications at three different levels.

First, from the very beginning of this study, ANT has performed as a meta-theoretical lens for the study design by suggesting consideration of the lived experiences and perceptions layer, the organizational layer and the technological layer as researchers interact with open access repositories for their knowledge production. The choice of socio-technological approach also informed the data collection methods by including interviews with the researchers as well as the use of primary documentary evidence with respect to the organizational and technological properties of the access tools and the open access repositories. As the findings reflect, actors and properties situated in all three layers have had implications on the co-construction dynamics that emerge as a result of the availability of the open access repositories to scholars in two distinct

disciplines: astronomy and philosophy of science. For example, the emergence of trust as one of the properties of ADS has been co-constructed over time. The technical integration capabilities of ADS (with arXiv and other systems and services, in the form of interfaces and interoperability standards) have been constructed by building the software to enable such capabilities, by using open standards and protocols. The organizers and the managers of ADS had to make a decision to build technical capabilities, with the researchers being one of the actors that perhaps provided feedback in terms of requested technical enhancements. The technical features then had to be instantiated into usable organizational level, an example of which is “Find Similar Abstracts”. By integrating and including articles and abstracts from multitude of sources, over time ADS has been perceived to be more and more inclusive and integrated. Thus, it emerges as being trusted by the astronomers to provide an almost complete view of the disciplinary knowledge ecosystem, in essence performing as a centralized proxy to the disciplinary knowledge ecosystem that can be accessed through a normalized interface. The trust has then performed on the tenure and proportion committees to use ADS as a citation impact assessment tool.

Second, tracing the relationships amongst the actors of any one of the layers and also between the three layers shows that actors from the lived experiences layer, the technological layer, and the organizational layer perform on each other by realigning the enacted processes. The realignment occurred at a very granular level such as the mutually co-constructed technological interoperability between arXiv and ADS, as well as the mutual co-construction at the epistemic level between individual and the disciplinary cultures.

ANT has made it possible to traverse between the granular (more concrete) and the broader context (more abstract) as it has been demonstrated in the previous example. Each step of the way, in each of the layers, multitude of actors have emerged and their properties have been identified and interpreted in relation to the four themes, easily crossing the boundaries between the technological and non-technological actors. Interpreting the specifics of how multitude of actors informed each other was enabled by ANT's inscription and translation concepts. An example of inscription would be the transposition of the paper based open collaboration culture onto the electronic scholarly communication; it can be described as the open collaboration's property of "openness" being inscribed onto the electronic scholarly communication context, and in the process reconfiguring it into open access information environment, where for example the postal mail for exchange has been translated into electronic means of communication. Within each of the layers and between the layers, actors perform on each other by realigning their respective contexts. The nature of realignment has emerged to be of two types: restructuring and emergence of new structures. Also, the openness concept as instantiated in practice and concrete steps by the open access repositories (artifacts being openly exchanged), has been translated and inscribed onto a disciplinary level context (open entrance for researchers into disciplinary discourse), thus open access has triggered realignment in the scholarly communication process to allow for new entrants and potentially realign the disciplinary knowledge ecosystem as an outcome.

ANT's approach of understanding and relating actors to each other through their properties is also congruent with the grounded theory methods for data analysis and interpretation used in this study. The axial coding specifically relies on using properties

to identify the emerging concepts, categories and themes. The congruency between the grounded theory and the actor-network theory enables a tracing capability between the organizational, technological and the perceptions layers. This tracing provides the investigator with a lens to see patterns across the different layers, and relationships that link technological and non-technological actors. For example, the emergence of the four themes relies on relationships in the perceptions layer (about the explicit or implicit relationships between scholarly process and scholarly output as stated by the researchers during the interview), with implications from the organizational and technological layers (about the visible or the invisible organizational and technological features and capabilities). The emergence and structuring of the individual knowledge production context construct, that has shown to be a valuable construct in this study, exhibits elements of process and the tools that enable the process, knowledge network and relationships with organizational structure.

Third, instead of analyzing the technological and organizational properties of various actors separately from each other and then analyzing the high level findings to understand the relationships between the technological and the social, ANT has enabled the socio-technological relationship to be identified and analyzed at granular level. Thus, instead of asking research questions from either the social or the technological deterministic perspective, thus removing the plausibility of any relationships between the social and the technological, ANT provides the investigator to be able to see whether there are any socio-technological relationships. If for example the building of the software capabilities of ADS that have enabled open interoperability is analyzed as technical task only, it would be hard to understand and trace why and how it was

implemented. It could have been a political decision or a funding based decision. Understanding this aspect might be helpful in understanding the scope of interoperability and inclusion of the various sources.

ANT has also provided a co-constructionist view into the interplay of actors from the three different layers: lived experiences and perceptions, organizational and technological. Whereas Fry's (2006) findings present a social-deterministic approach to the construction and use of digital resources by specifying disciplinary norms and cultures as actors that perform upon the scholarly practices of researchers, ANT has enabled the observation of a co-constructionist dynamic whereas the open access repositories and access tool have certainly been informed by disciplinary culture, but they have concurrently informed aspects of the disciplinary culture that has appropriated the open access repositories and the access tools as actors into the disciplinary knowledge production context.

## **Chapter 14. Evaluation criteria**

In this chapter, the researcher self-assesses the quality of the grounded theory based study, with respect to the research process and the grounding of the study in the data. The evaluation is guided by Strauss and Corbin (1998) and Creswell (1998) who present the following two sets of criteria for evaluating the quality and fitness of a qualitative study: a) criteria judging the adequacy of the research process, and b) criteria for judging the empirical grounding of a study. The goal is to help the reader understand the research process steps of the study and to provide guidance about how the findings of this study are grounded in the data.

Strauss and Corbin (1998, p. 269) list the following seven criteria for judging the adequacy of the research process. These criteria have been used by other studies that have used grounded theory approach and are also advanced by Creswell (1998, pp. 209-10). Each of these criteria is addressed and explained how this study meets them.

- Criterion 1: How was the original sample collected? On what grounds?

Selection of the OA repositories was carried purposefully to identify sciences and humanities OA repositories to introduce variations in the study, following the domain-analytic approach, in order to identify possible patterns of dynamics between the two distinct disciplinary cultures.

The participants were selected purposefully to meet the main criteria that they are using open access repositories in their knowledge production process with the intent to use the scholarly artifact found therein in the production of knowledge artifacts.

- Criterion 2: What major categories emerged?

Four main themes of discourse emerged as perceived by the researchers in their interaction with the open access resources. Researchers perceive that open access resources are implicated in their knowledge production process and the structuring of their local and individual knowledge network, are integrated with the scholarly context, and also have democratized the scholarly communication context.

- Criterion 3: What were some of the events, incidents, or actions (indicators) that pointed to some of these major categories?

For the astronomers, arXiv and ADS are perceived as very integrated with the disciplinary knowledge production context and very central in how scholar produced scholarly knowledge artifacts. The philosophers of science however do not necessarily view PhilSci as very central and integrate in their disciplinary culture.

- Criterion 4: On the basis of what categories did theoretical sampling proceed?

That is, how did theoretical formulations guide some of the data collection? After the theoretical sampling was done, how representative of the data did the categories prove to be?

The four major themes are inclusive from the perspective that they encompass the rest of the categories that emerged from the open and axial coding, such as research processes, the repositories, roles and values, access tools, and data repositories.



- Criterion 5: What were some of the hypotheses pertaining to conceptual relations (i.e., among categories), and on what grounds were they formulated and validated?

Some of the high level hypothesis were guided by the actor-network theory that posits that social and organizational level actors are interrelated and ought to be analyzed together in order to understand and interpret the socio-technological dynamic that have implication on researchers knowledge production processes and individual knowledge networks.

The findings show that ANT's methodological and theoretical guidance has enabled to understand and describe, more clearly and in a greater detail, the socio-technological dynamics as they are constructed between the organizational and the technological layers and how these have impacted the knowledge production processes.

- Criterion 6: Were there instances in which hypotheses did not explain what was happening in the data? How were these discrepancies accounted for? Were hypotheses modified?

The high-level hypothesis to examine the socio-technological dynamics of the scholarly context provided guidance for data collection. The data collection relies on interviews with researchers and documentary evidence that describes the organizational and technological layers of the open access repositories and the access tools. The difference about why arXiv scientists perceive the arXiv repository to be so strongly integrated with their disciplinary culture, and why PhilSci scholars do not perceive strong role of the PhilSci repository in their

disciplinary culture, may be explained by understanding the properties of the respective disciplinary cultures. An existing theoretical framework that shows differences between disciplinary cultures has been used to aid in the explanation.

- Criterion 7: How and why was the core category selected? Was this collection sudden or gradual, and was it difficult or easy? On what grounds were the final analytic decisions made?

The study proceeded with the four themes as guiding categories for the study, because they encompass the other categories that emerged from the initial open and axial coding, and because these four categories had the potential to relate researchers' individual production processes with the disciplinary production context. The four categories emerged gradually after a number of iterations between the open and the axial coding.

Similarly to the criteria for assessing the adequacy of the research process, Strauss and Corbin (1998, p. 270-2) list the following criteria for judging the empirical grounding of the study. These eight criteria are addressed to judge the preliminary empirical grounding of this study. These criteria have been already addressed partially by accounting and presenting a congruent theoretical, methodological and methods approach, and a congruent set of research questions.

- Criterion 1: Are concepts generated?

Four main concepts were generated, expressed as themes of discourse and are inclusive of the rest of the concepts. Further, a number of defining properties

emerged for each of the four concepts (i.e., themes). The defining properties emerged to be perceived differently between the two groups of researchers.

- Criterion 2: Are the concepts systematically related?

The concepts represented by the four themes are systematically related across the lived experiences and perception layer, and the organizational and technological layers. The two groups of researchers differ in their perception of some of the concepts. These differences largely are explained based on the data and the findings of this study; the rest are explained by relating to existing theoretical frameworks and findings in the information science scholarly literature.

- Criterion 3: Are there many conceptual linkages, and are the categories well developed? Do categories have conceptual density?

The four main themes are high-level contexts that in themselves are constructed entities with actors from the lived experiences and perceptions, organizational and technological layers. Each theme has been developed and explicated through its properties across the three different layers. Some concepts such as openness, time, and space, are common across the four themes. Other concepts are specific or more defining of a particular theme.

- Criterion 4: Is variation built into the theory?

Variation is build into the theoretical approach by selecting two distinct scholarly disciplines, humanities and sciences. These variations are evident in the study findings and the resultant propositions.

- Criterion 5: Are the conditions under which variation can be found built into the study and explained?

The differences between the two groups of researchers are explained and interpreted in the appropriate sections.

- Criterion 6: Has process been taken into account?

Process has been an integral part and a driving force of this study. The participants were interviewed about their knowledge production process, and the documentary evidence is intended to further explain the knowledge work of the participants.

- Criterion 7: Do the theoretical findings seem significant, and to what extent?

The findings seem significant at contextual, theoretical and methodological level. This study has contributed to the understanding about the role of open access in scholarly communication. Theoretically, it has contributed to the domain-analytic approach to understanding the differences in information practices amongst two disciplines. Methodologically, this study has demonstrated that the actor-network guided co-construction approach in conjunction with grounded-theory is a viable methodological approach for understanding and interpreting the role of complex socio-technological contexts used by users.

- Criterion 8: Does the theory stand the test of time and become part of the discussion and ideas exchanged among relevant social and professional groups?
- The outcomes of this study are structured in the form of findings and propositions for further study.

## **Chapter 15. Conclusion**

This study provides an understanding about the implication of OA repositories in the scholarly knowledge production process. Informed by studies in LIS showing that humanities scholars and scientists enact different information practices in their scholarly knowledge production, two groups of participants were selected in these scholarly disciplines. From the sciences disciplines, the recruitment process resulted in a group of astronomers and astrophysicists (as users of the arXiv OA repository), representing a discipline with high mutual dependence and low task uncertainty (see Whitley 2000; Fry 2006). From the humanities discipline, the recruitment process resulted in a group of philosophers of sciences (as users of the PhilSci OA repository), representing a discipline with low mutual dependence and high task uncertainty. The analysis and interpretation of the interviews with the researchers, guided by the grounded theory, provides an insight into their lived experiences and perceptions. The researchers perceive that OA repositories have implications for their knowledge production process, their knowledge output, are integrated with the broader scholarly context, and are able to democratize the scholarly discourse. Primary documentary evidence describing the open access repositories and the access tools was collected, guided by the actor-network theory and the socio-technological manifestation of OA repositories and related access tools. The analysis and interpretation of the documentary evidence reveals that actors from the technological and organizational layers of the open access repositories and the access tools perform on each other and are co-constructed within the broader disciplinary context.

### *Lived experiences and perception level*

The findings show that for astronomers the open access repository arXiv is centrally integrated in their discipline and trusted by the scientists individually on a daily basis to provide a complete view of the disciplinary knowledge ecosystem, either directly or through the use of ADS, as witnessed by all six scientists that participated in this study. For the philosophers of science, the open access repository PhilSci is perceived as only one of the ways through which they discover open access resources, and it is only marginally relevant for the PhilSci scholars in their daily scholarly lives. By and large, the Philosophers of science augment their search process by scholars' Home Pages and invisible colleges.

### *Organizational level*

The concept of knowledge production context has been applied in this study as a construct that contains the context denoted by the knowledge production process, the knowledge network and the relevant disciplinary norms and cultures. From this perspective, open access repositories are co-constructed over time within the broader context of the disciplinary knowledge production contexts. Thus, within the context of disciplines' institutionalization, they are mutually performed with the individual knowledge production contexts that are in turn situated and co-constructed within the disciplinary culture. Researchers' individual knowledge production contexts are interrelated and co-constructed as researchers enact their information practices as they go about searching for scholarly artifacts to be used in their knowledge production. The information practices are enacted within the boundaries of the possible feature and

functional capabilities made available by the technological layers of the open access resources and as appropriated and adapted by the organizing structures of the open access resources. Concurrently and over a period of time, researchers may need and may introduce new sets of tools and capabilities to make their information practices more efficient and more productive. Thus, the information practices of researchers are co-constructed with the open access resources over time. The feedback mechanism for the astronomers is open and receptive. The organizational and the technological layers of arXiv and ADS are co-constructed within the disciplinary boundaries. For the philosophers of science the feedback mechanism is disconnected as EPrints is developed outside of their disciplinary boundaries. Further, over time the information practices and the open access resources together may be co-constructed with disciplinary cultural norms and practices as they may explicitly or implicitly become accepted and standardized disciplinary norms and practices.

### *Technological level*

The analysis also shows that the technological layers enable almost the same features and functional capabilities for both groups of researchers for their interaction with the respective repositories. In addition, the organizing structures of arXiv and PhilSci have stated very similar intents and goals about the intended values that the repositories can bring to the scholarly community. This would suggest that both disciplines are positioned to benefit from the values provided by open access. Yet, arXiv has been adopted with a central and important role by the arXiv scientists, while PhilSci has been minimally appropriated by the PhilSci scholars. The demographic characteristics of the

researchers in the two disciplines do not explain why scientists perceive stronger implication of the OA repositories in comparison to the philosophers of science. The astronomers mostly seem further ahead in their career stages in comparison to the philosophers of science. Thus, the generation gap does not explain the difference; it actually might point in the opposite direction.

The dynamics between the technological, organizational and lived experiences layers as explicated in this study using ANT do not fully explain the difference about the role the open access repositories play within their respective disciplines. However, the strong co-construction dynamics suggest a viable common construct with conditions and properties that perform both on the scholarly individual knowledge production contexts and on the open access repositories. One such element or a set of elements, co-constructed as constitutive actors in the disciplinary knowledge production context, has been identified by Fry (2006), and together with the findings from this study has been suggested as propositions P2 and P3.

This study also demonstrates that the socio-technological co-constructionist approach, informed by the actor-network and by the grounded theory, is a viable methodological approach for understanding, describing and interpreting the relations between lived experiences and the socio-technological context where lived experiences are enacted.



### *15.1 Implications and study importance*

Addressing the implications of the open access phenomenon for the scholarly communication through the lived experiences of two distinct and different groups of researchers, astronomers and philosophers of science, as the findings have shown, has implications at two different levels. Contextually, this study has explicated that the socio-technological dynamics are implicated in the co-construction of disciplinary knowledge production contexts, through the interaction dynamics that are enacted as humanities scholars and scientists search, discover and access knowledge artifacts in open access repositories. Methodologically, the study has shown that the combination of the actor-network theory and grounded theory is a viable theoretical and methodological framework to analyze and interpret interaction dynamics that exhibit socio-technological properties.

#### *15.1.1 Contribution to LIS and scholarly communication*

Overall, the study has shown that open access repositories and access tools are perceived by researchers to be important actors in scholarly communication, with disciplinary characteristics and properties impacting the degree of importance. As socio-technological constructs, open access repositories are perceived to have impact on researchers' knowledge production processes and their individual knowledge networks, as well as on the broader disciplinary context, such as the democratization of the scholarly discourse.

Through the analysis and the interpretation of the interviews and documentary analysis, the study has identified numerous actors and properties attributed to open access repositories and the access tool used by the researchers, summarized in Table II

and Table I2. Some of these actors and their properties have been used in this study from a qualitative perspective. The tables are matrices of perceived properties of the access tools and open access repositories, and their relationship with the values as perceived by the researchers with respect to each theme.

As it has been described in the summary of findings, the meaning of a “property” (i.e., the term used to describe an actor) has different connotations at the technological, organizational, and lived experiences and perceptions layers. At the technological layer, the properties of the access tools and the repositories are features and functional capabilities provided by the software (such as EPrints), as built by the software developers. This is what the software is capable off. At the organizational layer, the properties are intended goals and values as instantiated by the organizers and managers the open access repositories (such as arXiv, PhilSci) and related policies, with the intent to be used by the scholarly community. Thus, an analysis related to properties from different layers needs to account for the fact that the properties are not necessarily intrinsic and basic; rather, they are also co-constructed.

The implication of the disciplinary cultures or knowledge domains on researchers’ knowledge work and information practices has been addressed by Bates (1994), Palmer (2005), Palmer and Cragin (2008), Whitley (2000) and Fry (2006). The underlying pattern of differences across the disciplinary fields is informed (either explicitly or implicitly) by the domain-analytic approach for understanding the commonalities but especially the differences in information practices between researchers in humanities and sciences. The domain-analytic approach in Information Science (IS) has been

carefully articulated and related to other approaches in the field of IS and has been advanced by Hjørland and Albrechtsen's (1995)

For IS our main thesis is that the point of departure is knowledge-domains, disciplines, or trades, not individuals and especially not the more biological, physiological, and psychological make-ups of individuals. The individuals should be seen as members of working groups, disciplines, thought or discourse communities, etc. IS should in other words be seen as a social science rather than as a cognitive science. (Cognitive science here understood in the mentalistic, intrapsychical tradition, not in the sociocognitive meaning.) (p. 409)

This study also contributes to the domain-analysis approach by building on Fry's (2006) findings. To summarize, Fry (2006) has found that disciplinary communities that exhibit high degree of mutual dependence in conjunction with low degree of research problem uncertainty (such as astronomy and astrophysics), are more likely to utilize and co-construct digital tools and resources. And, disciplinary communities that exhibit low degree of mutual dependence in conjunction with high degree of research problem uncertainty (such as philosophy of science), are less likely to utilize and co-construct digital tools and resources to the fullest potential (p. 299). As it has been shown in this study, open access has emerged as one of the actors that impacts the co-construction dynamics of researchers' information practices, where the implications of open access differ between the two disciplinary cultures that exhibit the typological characteristics in Fry's study.

Whitley's (2000) typology for grouping scholarly disciplines is better suited to categorize disciplines that exhibit strong and unambiguous patterns of mutual

dependence and task uncertainty, and cannot clearly account for the wide range of non-sciences scholarly disciplines where the researchers are using information technology in their information work. For example, using Whitley's approach, the discipline of classical studies will be categorized as a discipline with low mutual dependence and high task uncertainty. However, as Ruhleder (1994) has shown, the classical studies have benefited from the information technology: "The construction of the TLG [Thesaurus Linguae Graecae, an online textual database] has helped further stabilize, preserve, and disseminate a far wider range of texts than any single library has been able to collect to date, just as the printing press did five centuries earlier" (p. 224). This anomalous behavior of Whitley's approach is not addressed in this study. As a future study, it will be interesting to investigate the role of OA in disciplines exhibiting the disciplinary characteristics of classical studies.

Savolainen (2007) further compares and contrasts the value and the implications of the "information behavior" and "information practice" as umbrella concepts for research in IS, and specifically emphasizes the congruency between the domain-analytic approach and the study of information practices: "the discourse on information practice has affinities with the domain analytic approach ... For example, in the study of scholars' information practices, it is important to discuss the ways in which these practices are embedded within the overarching context of disciplinary differences in order to develop a holistic understanding of scholarly communities' work and communication practices" (p. 120).

With respect to the three meta-theories that inform IS, explicated in great detail by Talja, Tuominen, and Savolainen (2005), this study is informed by the "Collectivism"

(or “Social Constructivism”) meta-theory that is most closely exemplified by the domain-analytic approach (p. 82).

#### 15.1.2 Contribution to theory and methodology

In addition to the contextual contribution related to the implications of OA for scholarly communication, this study shows that the socio-technological co-constructionist theoretical approach guided by actor-network theory, in conjunction with grounded theory methods is a viable methodological approach for describing, interpreting and understanding relationships between lived experiences and the socio-technological context within which researchers’ lived experiences are enacted. To this extent, the theoretical and methodological approach of this study may be used to understand other contexts where users access socio-technological resources through intermediaries such as digital access tools. In order to transplant this theoretical and methodological approach to address other contexts, the level of congruency between the context of this study and other potential context needs to be assessed. For example, the context of this study is about researchers’ interactions with the aim to discover and access articles from open access repositories. If however the new context and the research problem is more about performing certain set of tasks at the systems that are being accessed and used by the users, different methods for data collection may be applied. The choice of data collection methods will have implications for the data analysis and interpretation steps as well, as it has been shown in section 6.1. The meta-theoretical co-constructionist approach of this study can also guide studies of different contexts by providing the research framework that emphasizes on users’ experiences and their use of socio-technological artifacts, with the ultimate aim to enable researchers to

bridge and analyze the social and the technological together, as has been shown in this study. This will bring forth relevant relationships between the social and the technological context that might be hidden in plain sight if the social context and the technological context of a phenomenon are studied separately.

### 15.1.3 Contribution to practice

The findings of the differences in the co-construction of open access resources at individual level and the co-construction of an intellectual field's social and epistemic norms and practices between intellectual fields in two distinct disciplines suggests that the same type of digital resources are integrated differently into researchers' individual scholarly practices. Considering that digital resources have been utilized across different disciplines and their intellectual fields, digital resources that are most congruent with the intellectual field should be developed, and those are unique to specific disciplines or to groups of intellectual fields that have common characteristics based on Whitley's (2000) "mutual dependence" and "task uncertainty" properties.

So, for example, for the astronomers the astro-ph section of arXiv and ADS are co-constructed with some of the epistemic and social properties of the intellectual field, i.e., astronomy and astrophysics. The construction of astro-ph and ADS as trusted, integral, interconnected and central digital resources is congruent with Fry's (2006) findings, where astronomy and astrophysics exhibit high level of mutual dependence and low level of task uncertainty. More specifically, Whitley (2000) points out that

Increases in the degree of functional dependence are associated with greater specialization of research topics and tasks, standardization of work procedures, competence standards and communication structure, and co-ordination of task

outcomes from different research sites for dealing with particular problems. The scope of problems tackled by individuals and research groups tends to decline as functional dependency grows. (p. 94)

This is congruent with standardized tasks and procedures and communication structures that have been enacted by the astronomers in their localized everyday scholarly work. The high degree of standardization and routines that can be quite complex, means that they can be delegated and inscribed into information technology systems and automated to great extent. The central nature of arXiv and ADS are certainly global and serve the discipline, but they are also very local from the perspective that these resources are used by the arXiv scientist to perform local tasks.

However, for the philosophers of science, the task uncertainty seems to be very high, especially the technical task uncertainty:

Generally, growing technical task uncertainty implies greater reliance on upon direct and personal control of how research is carried out, considerable local variations in work goals and processes and more informal communication and co-ordination processes. (Whitley 2000, p. 131)

This is congruent with the finding in this study that the PhilSci researchers were the point of collection and co-ordination of both epistemic and technical processes, whereas arXiv scientist had delegated and inscribed many of their procedural and routines tasks to arXiv and ADS.

Thus, the open access repositories and the access tools have distinctly different roles in these two disciplines. The appropriation, adoption and emergence of arXiv and ADS within the scholarly community of astronomers and astrophysicists are congruent with

the disciplinary norms and practices. For the philosophers of science, the structure of the repositories and the access tools, is not perceived as valuable (and therefore not very strongly co-constructed with the discipline itself), perhaps because philosophers of science might benefit more from a technology that will enhance their ability to personally reach more variety of resources and need to enact processes that are not easily inscribable and automatable into information technology for automated enactment. Philosophers of science may benefit more from usable online annotation tools with features such as recording and recalling different resources, tracking the action across the different sources, etc.

### *15.2 Study scope and limitations*

The findings of this study may not be generalized beyond the specific groups of researchers, except that they may be related to information practices in disciplinary fields or knowledge domains. One of the participant selection criteria in this study is that the participants use materials from open access repositories in their knowledge production process. The groups of researchers that accepted to participate in this study can thus be considered early adopters. However, as one of the findings suggested that the arXiv participants perceive the arXiv open access repository as integral and central in their discipline, this might suggest that the findings may be applicable more generally to other users of astro-ph, which was the specific sub-section of arXiv used by the arXiv scientists that participated in this study. The total number of participants, eleven, does not provide for sound generalizability, however it is adequate for qualitative study whose goal is to understand and explicate a new phenomenon as perceived by a group of purposefully selected people because they experience the same phenomenon.



The study is also limited to participants using resources found in the open access repositories, and it does not extend to the value the participants see in submitting contributions to the repositories. The task of depositing the articles by the authors is out of scope of this study and therefore not addressed as part of the interaction. Although some of the participants reflected on the values that the repositories provide for the distribution of their own work.

In the literature review, the Copyright Transfer Agreements (CTAs) emerged as possible actors in researchers' interaction with open access repositories. However, neither during the interviews with the researchers, nor from the documentary evidence, the CTAs emerge as relevant actors that impact how researchers use materials found in the open access repositories. Thus, this study does not reveal much about the role of CTAs, beyond what has been learned from the literature review.

This study does not cover the cognitive aspects of researchers' information seeking process. Rather it addresses researchers' information practices as defined by Savolainen (2007)

... a basic characteristic of the discourse on practice, in general, as well as "information practice," in particular, is the emphasis placed on the role of contextual factors of information seeking, use, and sharing, as distinct from the individualist and often decontextualized approaches that are seen as characteristic of assumptions of information behavior. (p. 121)

### *15.3 Propositions*

Next, the main findings of this study are structured into propositions that can be further studied as either a part of quantitative, qualitative or mix methods studies. Also,

some of the additional questions that emerged from the interpretations in this study but could not be explicated based on the collected data are formulated into propositions for further research.

Reflecting on Table 26 that shows the perceived difference by both groups of researchers with respect to the four themes, the difference between the two groups of researchers can be stated as a proposition that relates the individual knowledge production contexts with the disciplinary knowledge production contexts:

P1a: Open access repositories that are perceived by the researchers to have strong implication on their scholarly process and scholarly output are also perceived to be strongly integrated with their disciplinary context and contribute to the democratization of scholarly discourse.

Alternatively, from the perspective of individual researcher's and disciplinary knowledge production contexts, the proposition may be stated as:

P1b: Open access repositories that are perceived to have strong implication on researcher's individual knowledge production contexts are also perceived to be strongly integrated with their disciplinary knowledge production context.

Reflecting on the difference about the development and emergence of specialized access tools such as ADS, in relation to a specific open access repository, the answer needs to be addressed by consulting additional literature that describes properties of the epistemic cultures that are beyond the scope of this study. As it has been already shown, one such element or a set of elements, co-constructed as constitutive actors in the disciplinary epistemic culture, has been identified by Fry (2006). In her study about the

relationships of disciplinary social and epistemic cultures with how communication channels and digital resources are utilized by the members of a specific discipline, Fry finds that disciplinary communities that exhibit high degree of mutual dependence in conjunction with low degree of research problem uncertainty (such as astronomy and astrophysics), are more likely to utilize and co-construct digital tools and resources that will make their information practices more productive and more efficient. Similarly, disciplinary communities that exhibit low degree of mutual dependence in conjunction with high degree of research problem uncertainty (such as philosophy of science), are less likely to utilize and co-construct digital tools and resources to the fullest potential (p. 299).

This finding by Fry (2006) is advanced by the findings in this study, thus resulting in the following propositions for further consideration and research:

P2: In a given discipline, the roles and values of an OA repository and the associated tools are strongly affected by the degree of mutual dependence amongst the researchers and the level of task uncertainty in dealing with the research problems.

P3: The higher the mutual dependence amongst the researchers of a discipline, in conjunction with low level of task uncertainty, the more likely the OA repository will be strongly integrated within the disciplinary culture and the routine activities of researchers.

And a corollary proposition,

P4: Specialized access tools for open access repositories will emerge in disciplines that exhibit high degree of mutual dependence and low level of task uncertainty.

#### *15.4 Further research*

##### *Same context*

To understand further the role of OA in the scholarly communication and its implication for knowledge production, the qualitative findings of this study can be used to develop hypotheses to test the propositions and findings of this study. Specific to the disciplines represented by the participants of this study, one approach will be to distribute a survey to the two discussion lists that were identified as part of the participant recruitment process. The data collection instrument, a survey in this case, will be designed using the qualitative findings of this study, more specifically the various open access properties that are central in defining each theme and various aspects of researchers information practices. A quantitative study can also be expanded beyond these two disciplines, to include disciplines in different stages of their open access repository implementations in relation to their mutual dependence and task uncertainty, using Whitley's (2000) theory.

It would be also interesting to identify the actors and their properties that can explain the emergence or non-emergence and adoption of trusted and centrally positioned open access resources such as arXiv and ADS in fields that do exhibit similar characteristics in terms of mutual dependence and task uncertainty such as astronomy and astrophysics. With respect to this, Fry has already identified the following next step: "Further work needs to be done, therefore, to develop a systematic operationalization of Whitley's theory across a diverse range of case studies and to construct both qualitative and quantitative indicators for 'mutual dependence' and 'task uncertainty'" (Fry 2006, p. 313). This further work suggested by Fry can be enhanced by explicit consideration of

open access as one of the actors that can affect “the qualitative and quantitative indicators.”

#### Different context - transferability

Next, the socio-technological co-constructionist approach that provides the theoretical and methodological framework for this study can be assessed for applicability to analyze and explicate similar context that exhibit socio-technological manifestation. Of special interest would be contexts where users interact with information environments through intermediaries, with the goal of using the information environment to create new knowledge or enhance innovation. The collaborative but loosely coupled Wikipedia material construction may be one such context. Should the collaborative production of Wikipedia materials be studied as “open access” (the approach used in this study), “open source” (the framework used to understand the dynamics in the production of open source software) or “open content” (the framework of the production of media content)? All three have at least one common element, the openness of the “thing” that is being produced: knowledge artifacts being made “free”, free software (a type of procedural and algorithmic knowledge), and open or “free” content (mostly block of information as well as images). In order to apply the approach used in this study to the collaborative efforts in the production of Wikipedia materials, the congruency between the concept of open access and open content need to be assessed, as well as the congruency of the collaboration dynamics between the producers of the content.

### Methodological research

As a long-term thinking, if the methodological approach used in this study emerges as a congruent approach for understanding and analyzing socio-technological contexts, it will be beneficial to extract the methodological approach out of the context and build it into a more inclusive (with respect to context and type of users) approach that may be used to analyze other user experiences. A key element of such a study will be to determine the level of applicability and congruency of this approach by assessing the variations in the lived experience contexts, the variations of intermediaries used, and the variations with respect to the resources being accessed.

## Appendix A: Types of research studies about OA

citation	pub date	quantitative	qualitative	OA access flavor	article level	journal level	perception / behavior / attitude	citation / bibliog.	use / log analysis	usability	OA impact on the knowledge production process
Wilson (2005)	2005		x	institutional repository	x		x		x		
Antelman (2004)	2004	x	x	self archiving	x			x			
Nicholas, Huntington and Rawlands (2005)	2005	x	x	various OA means	x		x				
Hoorn & van der Graaf (2006)	2006	x	x	various OA means	x		x				
Schauder (1994)	1994	x	x		x	x					
Pavlisca (1996)	1996		x	mix OA and e-j		x	x			x	
Speier et al. (1999)	1999	x	x	mix OA and e-j		x	x				
Hedlund, Gustafsson and Björk (2004)	2004	x	x	various OA means		x		x		x	
Harnad (May 2006)	2006	x		institutional repository	x			x	x		
Rowlands & Nicholas (2005)	2005	x		institutional repository	x		x				
Phanouriou et al. (1999)	1999	x		institutional repository	x					x	
Creaser (2006, May)	2006	x		mixed (sub/OA)	x			x	x		
Thelwall and Harries (2003)	2003	x		personal home page	x			x			
Brown (2001)	2001	x		preprints	x		x	x	x		

<b>citation</b>	<b>pub date</b>	<b>quantitative</b>	<b>qualitative</b>	<b>OA access flavor</b>	<b>article level</b>	<b>journal level</b>	<b>perception / behavior / attitude</b>	<b>citation / bibliog.</b>	<b>use / log analysis</b>	<b>usability</b>	<b>OA impact on the knowledge production process</b>
Warr (2002)	2002	x		preprints	x			x	x		
Brown (2003)	2003	x		preprints	x		x	x			
Lawal (2002)	2002	x		preprints	x		x		x		
Nicholas et al (2006, June)	2006	x		pure OA	x				x		
Harnad and Brody (2004)	2004	x		self archiving	x			x			
Swan & Brown (2005)	2005	x		self archiving	x		x				
Swan & Brown (Feb 2004)	2004	x		various OA journals	x	x	x	x	x	x	
Kurtz et al. (2005)	2005	x		various OA means	x			x			
Rowlands and Nicholas (2005)	2005	x		various OA means	x	x	x				
Beier & Tschida (2003)	2003			various OA means	x	x					
Kurata et al. (2006)	2006	x			x		x			x	
Kortelainen (2004)	2004	x			x	x			x	x	
Prosser (2003)	2003				x	x					
Wulff & Nixon (2004)	2004	x		maybe OA		x	x	x	x		
Harter (1998)	1998	x		mix OA and e-j		x		x			
Chu (1999)	1999	x		mix OA and e-j		x	x		x		
Rich & Rabine (2001)	2001	x		mix OA and e-j		x	x		x		
Fosmire & Young (2000)	2000	x		mix OA and e-j		x			x		



<b>citation</b>	<b>pub date</b>	<b>quantitative</b>	<b>qualitative</b>	<b>OA access flavor</b>	<b>article level</b>	<b>journal level</b>	<b>perception / behavior / attitude</b>	<b>citation / bibliog.</b>	<b>use / log analysis</b>	<b>usability</b>	<b>OA impact on the knowledge production process</b>
Saxby (2006, June)	2006	x		pure OA		x	x				
Zhang (1998)	1998	x		pure OA		x		x			
Llewellyn, Pellack, and Shonrock (2002)	2002	x		pure OA		x		x	x		
Fosmire & Yu (2000)	2000	x		various OA journals		x		x			
Hawkins (2001)	2001	x		various OA journals		x		x			
Thomson ISI (2004)	2004	x		various OA journals		x		x			
McVeigh & Pringle (2005)	2005	x		various OA means		x		x			
Swan & Brown (2004)	2004	x		various OA means		x	x				
Barker & Tedd (1999)	1999			web magazine / portal			x		x	x	
Groote et al. (2005)	2005	x				x		x		x	
Sathe et al. (2002)	2002	x				x	x		x		
Liu (2005)	2005	x				x	x		x		
Siebenberg et al. (2004)	2004	x				x			x		

## Appendix B: OA repositories selection process

### *Registry of Open Access Repositories (ROAR)*

The ROAR website is located at: <http://roar.eprints.org>

Table B1: List of all 17 ROAR OA disciplinary (cross-institutional) repositories for the US

Name	Discipline	URL	Number of records	Content type	Software
PubMed Central	Life science / biomedical	<a href="http://www.pubmedcentral.gov">http://www.pubmedcentral.gov</a>	1073207	OA Articles, as well as OA Journals	Unknown
CiteSeer.PSU (CiteSeer.IST)	Scientific literature	<a href="http://citeseer.ist.psu.edu">http://citeseer.ist.psu.edu</a>	767558	Indexes and references to articles	CiteSeer
arXiv.org	physics, mathematics, related fields	<a href="http://arxiv.org">http://arxiv.org</a>	460118	Articles, pre-prints, other	EPrints, other
Social Science Research Network (SSRN)	Social science	<a href="http://www.ssrn.com/">http://www.ssrn.com/</a>	138546	Articles and other	Unknown
BEACON eSpace at Jet Propulsion Laboratory	Jet propulsion lab; science	<a href="http://trs-new.jpl.nasa.gov/dspace">http://trs-new.jpl.nasa.gov/dspace</a>	21611	Articles and other	DSpace
antbase.org	Ant species	<a href="http://antbase.org">http://antbase.org</a>	7823	Articles	Unknown
Archive of European Integration	European integration and unification	<a href="http://aei.pitt.edu">http://aei.pitt.edu</a>	6008	Articles, other research, official documents	EPrints
Texas Digital Library		<a href="http://repositories.tdl.org">http://repositories.tdl.org</a>	4511		DSpace
Digital Library of the Commons	international literature on the commons	<a href="http://dlc.dlib.indiana.edu">http://dlc.dlib.indiana.edu</a>	2072	Articles and other	EPrints
Woods Hole Open Access Server (WHOAS)	Marine biology oceanography	<a href="https://darchive.mblwhoilibrary.org">https://darchive.mblwhoilibrary.org</a>	1894	Articles and other	DSpace
Home - Alliance Digital Repository	The Colorado Alliance of Research Libraries	<a href="http://adr.coalliance.org/adrlib">http://adr.coalliance.org/adrlib</a>	1772	Articles and other	Fedora
PhilSci: Philosophy of Science Archive	Philosophy of science	<a href="http://philsci-archive.pitt.edu">http://philsci-archive.pitt.edu</a>	1530	Pre-prints	EPrints
DLIST: Digital Library of Information Science and Technology	Information science and technology	<a href="http://dlist.sir.arizona.edu">http://dlist.sir.arizona.edu</a>	1197	Articles and other	EPrints
Washington State University Research Exchange	Many disciplines	<a href="https://research.wsulibs.wsu.edu:8443/dspace">https://research.wsulibs.wsu.edu:8443/dspace</a>	850	Articles and other	DSpace

Name	Discipline	URL	Number of records	Content type	Software
Internet Archive		<a href="http://www.archive.org">http://www.archive.org</a>	563	Audio, video, text, etc.	Unknown
Aquatic Commons	Marine and aquatic science	<a href="http://aquacomm.fcla.edu">http://aquacomm.fcla.edu</a>	500	Articles	EPrints
CalTech: Fourth International Symposium on Cavitation		<a href="http://cav2001.library.caltech.edu">http://cav2001.library.caltech.edu</a>	111	Conference proceedings	EPrints

Note: Repositories were accessed on January 27, 2008. They were sorted by the descending number of records

*The Directory of Open Access Repositories (OpenDOAR)*

The OpenDOAR website is located at: <http://www.opendoar.org/>

Table B2: List of OpenDOAR disciplinary repositories for the US

Name	Discipline	URL	Number of records	Content type	Software
American Memory		<a href="http://memory.loc.gov/">http://memory.loc.gov/</a>	9000000	Articles	Unknown
PubMed Central		<a href="http://www.pubmedcentral.nih.gov/">http://www.pubmedcentral.nih.gov/</a>	800000	Articles	PMC
arXiv.org e-Print Archive		<a href="http://arxiv.org/">http://arxiv.org/</a>	500000	Article	arXiv
Research Papers in Economics		<a href="http://repec.org/">http://repec.org/</a>	302882	Articles	Unknown
LOUISiana Digital Library Server Repository		<a href="http://louisdl.louislibraries.org/">http://louisdl.louislibraries.org/</a>	36539	Articles	CONTENTdm
kydl OAI Archive		<a href="http://kdl.kyvl.org/">http://kdl.kyvl.org/</a>	36408	Articles	Unknown
AgEcon Search		<a href="http://agecon.lib.umn.edu/">http://agecon.lib.umn.edu/</a>	24799	Articles	Unknown
Digital Library for Earth System Education		<a href="http://www.dlese.org/library/">http://www.dlese.org/library/</a>	13362		Unknown
bepress Legal Repository		<a href="http://law.bepress.com/repository/">http://law.bepress.com/repository/</a>	10621	Articles	Bepress
Carlyle Letters Online: A Victorian Cultural Reference		<a href="http://carlyleletters.org/">http://carlyleletters.org/</a>	10000		Unknown
Archive of European Integration		<a href="http://aei.pitt.edu/">http://aei.pitt.edu/</a>	4803	Articles	EPrints
Open Video Project		<a href="http://www.open-video.org/">http://www.open-video.org/</a>	3967	Articles, theses	Unknown
Electronic Environmental Resources Library		<a href="http://www.eerl.org/">http://www.eerl.org/</a>	3814		Unknown
Center for Jewish History Digital		<a href="http://digital.cjh.org/">http://digital.cjh.org/</a>	3478	Articles	DigiTool

Name	Discipline	URL	Number of records	Content type	Software
Collections					
Northeast Massachusetts Digital Library: Imagining History Collections		<a href="http://nmrlsdli.cdmhost.com/">http://nmrlsdli.cdmhost.com/</a>	3149	Articles	CONTENTdm
Historic American Sheet Music		<a href="http://scriptorium.lib.duke.edu/sheetmusic/">http://scriptorium.lib.duke.edu/sheetmusic/</a>	3042		Unknown
Cold War International History Project		<a href="http://cwihip.si.edu/">http://cwihip.si.edu/</a>	2500		Unknown
RRUFF Project		<a href="http://rruff.geo.arizona.edu/rruff/">http://rruff.geo.arizona.edu/rruff/</a>	1733		Unknown
Central Florida Memory		<a href="http://www.cfmemory.org/">http://www.cfmemory.org/</a>	1701	Articles, theses	Unknown
Latin American Open Archives Portal		<a href="http://lanic.utexas.edu/project/laoap/">http://lanic.utexas.edu/project/laoap/</a>	1682	Articles	Unknown
Perseus Digital Library		<a href="http://www.perseus.tufts.edu/">http://www.perseus.tufts.edu/</a>	1591	Articles	Perseus

Note: Repositories were accessed on January 27, 2008. They were sorted by the descending number of records. The “Discipline” column is not populated as OpenDOAR did not capture this data field.

## **Appendix C: Request for Participation (RFP)**

This appendix contains the RFP text that is sent via e-mail and postal mail directly to the potential participants that have been selected.

### **Text of the RFP that will be sent directly to identified potential participants**

*Subject:* request to participate in a study that investigates scholars' interaction with self-archived open access repositories

Dear <insert actual name including academic title>,

I would like to request your cooperation in conducting this study that investigates scholars' interaction with self-archived open access repositories, using semi-structured interviews, document analysis and content analysis. This study aims to provide detailed understanding of scholars' interactions with open access repositories of self-archived peer-reviewed articles and other scholarly artifacts. These findings will contribute to research in library and information science and information systems design, and may be beneficial in providing a methodological approach for studying other socio-technological phenomena by the global scientific community.

#### **Objectives of the Study:**

The intent of this study is to provide an understanding of scholars' interactions with open access repositories of knowledge artifacts in the process of knowledge production, i.e., finding, discovering and accessing articles and other knowledge artifacts in OA repositories for use in knowledge production.

#### **Participant Characteristics:**

The participants will be scholars who are active users of OA repositories and use the OA repositories as resources to find, discover and access articles and other knowledge artifacts to be used in their knowledge production. Considering that you have published a number of articles in <insert name of repository>, I have identified you as an active user of OA repositories.

#### **The Interview Procedure:**

As a participant in this study, you will be asked to complete a short three question pre-interview questionnaire that also identifies your recently produced knowledge artifacts. The interview questions will be presented to you prior to the interview. The interview will last 60 minutes and it will be conducted in person or by phone by the investigator at a location and time mutually agreed by the investigator and the participant. The interview will be recorded by a hand-held digital recorder as well as laptop for backup. Access to the data will be available only to the investigator and a professional transcriber. There will be a post-interview discussion with the participant if there is a need for clarification around the interview data. There will be no physical, psychological, social, or legal risks involved in this process.

When?

The interviews will be scheduled between Nov 15<sup>th</sup>, 2008 and April 30<sup>th</sup>, 2009.

Outlets for distribution of research results:

The interview data will be analyzed in conjunction with document analysis and content analysis and the results will be reported in my doctoral dissertation. I also plan to publish several papers in library and information science and information systems journals. In reporting on the research, the confidentiality of the subjects will be assured. Any information obtained in connection with this study that can be identified with you will remain confidential and will be disclosed only with your permission.

If you would like to participate:

Please contact Mentor Cana at the address provided below. Once the time is set for the interview, you will receive the pre-interview questionnaire, the interview questions, and the consent form that includes information as mandated by the university. I will talk to you about the process and answer any questions you may have about the study.

Mentor Cana, Ph.D. Candidate

School of Communication, Information and Library Studies (SCILS)

Rutgers, The State University of New Jersey

4 Huntington Street

New Brunswick, New Jersey 08901

Tel: +1 (732) 690-2132

E-mail: [mentor@scils.rutgers.edu](mailto:mentor@scils.rutgers.edu)

## Appendix D: Informed Consent Form

### **Scholars' interaction with self-archived open access repositories: Applying actor-network theory to describe scholars' knowledge production**

I would like to request your cooperation in conducting this study that investigates scholars' interaction with self-archived open access repositories, using semi-structured interviews, document analysis and content analysis. The socio-technological co-constructionist approach provides the theoretical and methodological framework for this study. You were selected to participate in this study because as a scholar and an author you are an active user of open access repositories and you use open access repositories to find, discover and access knowledge artifacts which are used in your knowledge production. A total of sixteen participants are being recruited for this study.

This study aims to provide detailed understanding of scholars' interactions with open access repositories of self-archived peer-reviewed articles and other scholarly artifacts, through the discovery of systematic interaction patterns. These findings will contribute to research in library and information science and information systems design, and may be beneficial in providing a methodological approach for studying other socio-technological phenomena by the global scientific community.

If you should decide to participate, you will be asked to complete a short three question pre-interview questionnaire that also identifies your recently produced knowledge artifacts. The interview questions will be presented to you prior to the interview. The interview will last 60 minutes and it will be conducted in person by the investigator at a location mutually agreed by the investigator and the participant. The interview will be recorded by a hand-held digital recorder as well as laptop for backup. There will be a post-interview discussion if there is a need for clarification around the interview data. Possible risk factors from your participation are no greater than normal everyday activity. *Your decision to participate in this study is voluntary. Refusal to participate will involve no penalty or jeopardize your position.* You may withdraw at any time during the interview process and the recording and data taken up to the point that you withdraw will be destroyed if you request that I do so.

Any information obtained in connection with this study that can be identified with you will remain confidential and will be disclosed only with your permission. Only averages and other descriptive statistics will be reported in any publication. Only trained coders associated with the university will review the recordings and none will be able to identify you by name. In addition to using the collected data for this study, potential follow-up studies may be based partially or wholly on these data.

If you have any questions about your rights as a research subject, you may contact the IRB Administrator at Rutgers University at:

Rutgers University, the State University of New Jersey  
Institutional Review Board for the Protection of Human Subjects  
Office of Research and Sponsored Programs,  
3 Rutgers Plaza  
New Brunswick, NJ 08901-8559  
Tel: 732-932-0150 ext. 2104  
Email: humansubjects@orsp.rutgers.edu.

In case you have any questions related to the research study, the investigator may be reached at (732) 690-2132 or email: mentor@scils.rutgers.edu. The investigator's full address is:

Mentor Cana, Ph.D. Candidate  
School of Communication, Information and Library Studies  
Rutgers University, The State University of New Jersey  
4 Huntington Street  
New Brunswick, New Jersey 08901

---

My signature indicates that I have read the information above and have decided to participate. I realize that I may withdraw without prejudice at any time after signing this form should I decide to do so. If you desire a copy of this consent form, one will be provided for you.

Thank you.

Participant's signature \_\_\_\_\_ Date \_\_\_\_\_

Investigator's signature \_\_\_\_\_ Date \_\_\_\_\_



**Appendix E: Pre-interview e-mail questionnaire**

1: Would you please describe your general area of research?

2: Can you please describe the role of OA in your discipline/field?

3: Please identify up to five (5) knowledge artifacts (such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc.) you have recently produced that are a good representation of your use of the OA repository.

The interview questions will be related to these knowledge artifacts. You may identify the knowledge artifact that you believe is the best representation of your use of the OA repository.

*Instructions:* Please respond to these questions at least one week before the interview.

## Appendix F: Interview questions

1. Would you please introduce yourself with respect to your discipline and your use of OA repositories?
2. How would you describe the role of OA in scholarly communication?
3. How would you describe your involvement with the repository that you used in your work to produce knowledge artifacts such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc.?
4. How would you describe the production process of your knowledge artifact?
  - How did you go about finding and accessing relevant information and articles?
5. How would you assess the role of OA repositories in the different stages of your research and writing?
  - To what extent do you use OA repositories in the idea development stage?
  - How about later stages such as literature review or background information, argument building, interpretations, discussion, and conclusion?
6. How do you handle the information you discover in OA repositories?
7. What are the most frequent ways you discover the articles and other knowledge artifacts found in OA repositories?
8. What are some of the access tools you use to find information in OA repositories? Examples of access tools are: Google search, Google Scholar, Scirus, browsing of various indexes, browsing through topical reference databases, etc.
  - Do you have a preferred access tool? Why is this access tool preferred?
9. Can you easily identify some of the most relevant references you identified through OA repositories that used in your most recent work (such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc.)?
  - How important are these references that you have identified above in comparison to the references you have identified, located and/or accessed via non-OA venues?
  - Please describe the difference.
10. Would you please describe to what extent the institutional and disciplinary OA repositories you have used have been useful?
  - Can you please list some capabilities that you always use?
  - Can you please describe few cases/events in which OA repositories were extremely helpful in your research and writing?
11. Please list some of the challenges in using OA repositories.
12. What do you do when looking for an article you know it is out there, you have seen the references, but you cannot locate or cannot access it?
  - How are the OA repositories helpful in this case?
13. What are some of the improvements you would like to see with:
  - OA repositories
  - access tools
14. Do you have anything else you would like to add that we might not have had a chance to discuss yet?

## Appendix G: Interview Guide

- 1) Collect and fill the following data about the author before the interview:

Name:

Rank (assistant, associate, professor, etc):

Career stage:

Institution:

Discipline:

Age:

Gender:

Number of articles published so far:

Location of the interview:

Repository:

- 2) At least few days before the interview, read the article that the author has identified in the pre-interview questionnaire as his/her representative article about his/her use of OA repositories.
- 3) Before the start of the interview discuss with the author briefly the pre-interview response especially around the article that will be referred through the interview.
- 4) Proceed with the interview using the questions and the prompts listed below:

Interview questions and prompts:

1. Would you please introduce yourself with respect to your discipline and your use of OA repositories? (**descriptive, experience**)

*Prompts:*

- Further probe the specific relationships mentioned
- Further probe the specific use of the OA repositories
- Any specific role of OA in your discipline?

2. How would you describe the role of OA in scholarly communication? (**perception**)

*Prompts:*

- Is OA discipline dependent?

3. How would you describe your involvement with the repository [arXiv | PhilSci] that you used in your work to produce knowledge artifacts such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc.? **(process)**

*Prompts:*

- Relationships as an author

4. How would you describe the production process of your knowledge artifact? **(process)**

- How did you go about finding and accessing relevant information and articles?

*Prompts:*

- Describe some concrete steps
- OA or non-OA. Does it make difference?

5. How would you assess the role of OA repositories in the different stages of your research and writing? **(process)**

- To what extent do you use OA repositories in the idea development stage?
- How about later stages such as literature review or background information, argument building, interpretations, discussion, and conclusion?

*Prompts:*

- Positive role?
- Negative role?

6. How do you handle the information you discover in OA repositories? **(experience, process)**

*Prompts:*

- Desktop storage / indexing / categorizing / etc

7. What are the most frequent ways you discover the articles and other knowledge artifacts found in OA repositories? **(process, experience)**

8. What are some of the access tools you use to find information in OA repositories? Examples of access tools are: Google search, Google Scholar, Scirus, browsing of various indexes, browsing through topical reference databases, etc. **(factual, tech)**

- Do you have a preferred access tool? Why is this access tool preferred?  
(**experience**)

9. Can you easily identify some of the most relevant references you identified through OA repositories that used in your most recent work (such as articles, pre-prints, books, project reports, funding proposals, scholarly presentations, teaching materials, etc.)? (**factual**)

- How important are these references that you have identified above in comparison to the references you have identified, located and/or accessed via non-OA venues? (**experience**)
- Please describe the difference.

*Prompts:*

- General, your other knowledge artifacts.

10. Would you please describe to what extent the institutional and disciplinary OA repositories you have used have been useful? (**experience**)

- Can you please list some capabilities that you always use?
- Can you please describe few cases/events in which OA repositories were extremely helpful in your research and writing? (**experience, process**)

*Prompts:*

- Useful in the knowledge production process

11. Please list some of the challenges in using OA repositories. (**experience**)

12. What do you do when looking for an article you know it is out there, you have seen the references, but you cannot locate or cannot access it? (**process**)

- How are the OA repositories helpful in this case? (**process**)

13. What are some of the improvements you would like to see with: (**experience, descriptive**)

- OA repositories
- access tools

*Prompts:*

- Quality
- Quantity

14. Do you have anything else you would like to add that we might not have had a chance to discuss yet?

*Prompts:*

- Citation of pre-prints acceptable in your discipline?
- Meta-search across many OA repositories

## Appendix H: The emergent categories and the codes representing the four main themes

	Themes	
--	--------	--

Type	Name	Memo Link
	<b>Impact on scholarly process</b>	

Type	Name
	reduced distribution times
	collaborative work
	PhilSci, place to swap ideas
	collaborative env
	access to papers you can't find elsewhere
	open access is very critical
	arXiv speeds research
	arXiv increases visibility of papers
	more time for research
	PhilSci, ideas in making
	PhilSci speeds research
	articles found via open access or NOA equally important
	improved process and efficiency
	closed journals, barriers to knowledge discovery
	old materials
	daily updates for discovering ideas
	temporary collaboration space
	Home Pages very important
	conferences, PhilSci similar to
	PhilSci, scanning for recent papers
	cut out all that middleman
	PhilSci, a resource for latest research
	publication lag
	crucial for research
	essential to post in arXiv
	rarely visit library
	ADS used more than arXiv
	original material, maybe not published
	discover new things
	arXiv is huge

	vehicle for disseminating results
	rate of discussions accelerated
	important role
	ADS is location agnostic
	astro-ph increases article access
	journal backlog
	journal publication are slow
	PhilSci well done
	open access, delayed
	easier to do research
	connects people
	PhilSci, a preparatory place for final draft
	intellectual exchanges
	minor role
	purely open access repository is not essential
	wider readership
	articles found in JSTOR or PhilSci equally important
	couldn't do my research without open access
	improved search process
	improved research process
	open access can be a facilitator
	no need to visit journals directly any more
	absolutely essential

	<b>Impact on scholarly output</b>	
--	---------------------------------------	--

Type	Name
	access to things before you would have otherwise
	serendipity
	articles as triggers for new ideas
	cross-fertilization between disciplines
	subfields - a challenge to serendipity
	arXiv motivates new research
	arXiv increases citation rates
	don't reinvent the wheel
	PhilSci increases quality
	make ideas available early
	PhilSci might have important articles
	easier to share ideas
	PhilSci encourages research
	find latest research



	data from open access can motivate study
	open access makes possible research
	astro-ph increases quality
	huge difference
	being ahead of the time.
	open access increases quality
	PhilSci articles as important
	PhilSci enables to do more
	PhilSci, inspires new papers
	ADS and astro-ph provide the overall context
	Preprints are critical

	<b>Integration with scholarly context</b>	
--	---	--

Type	Name
	disciplinary culture
	PhilSci, not a publishing place
	arXiv, ADS, part of everyday life
	pre-print culture
	wider distribution
	arXiv helped career
	PhilSci, not as good as proper journal archive
	arXiv and ADS sufficient
	most commercial journal articles in arXiv
	everybody in the field knows about arXiv
	morning coffee
	PhilSci not a primary source
	PhilSci, source of upcoming articles
	no journal is missing from arXiv ADS
	arXiv use more than journal as a resource
	steel ideas easier
	open access moderately important
	open access just a facilitator
	open access has no prestige
	quality through peer review
	confident in using PhilSci materials
	PhilSci, the only archive used
	PhilSci, relied on it extensively
	PhilSci, is relevant
	PhilSci, is worth taking serious

	PhilSci, important because is disciplinary
	PhilSci, not peer-reviewed
	dislikes to use article from commercial journals
	journals still remain important
	invitation to speak
	not amazed any more
	open access, has changed communication

	<b>Democratization of the scholarly discourse</b>	
--	---	--

Type	Name
Tree Node	value for smaller institutions
Tree Node	democratic
Tree Node	small number of active researches in discipline
Tree Node	democratization of knowledge

## Excerpts from NVivo coding

dissertation

**Project:**

**Generated:**

4/15/2010 7:09 AM

**Name**

**Initials**

**Coding By**

MC

Mentor Cana

**Total Users**

1

**Internals\interviews\arXiv users\08121901\_a3**

**Document**

**References**

**Coverage**

2

**Tree Nodes\Themes\Democratization of the scholarly discourse\value for smaller institutions**

**Node Coding**

**2.87%**

*Reference*

*Character Range*

1

1.67%

Coverage

37972 -

38649

One thing I particularly like about this archive system is that it's extraordinarily democratic, that is that you don't have to be at a major research university to get access to it. When I went through my training and post-doc era, you had to be at a major research institute in order to be able to see the pre prints as people use to mail out through the airmail, and hard copies of papers to the major institutes in the world. So you typically send

out 300 or 400 of these copies of your article to all these different institutes, and if you were not at one of those institutes, you never got to see what was currently patenting in your area until the articles appeared.

#### Reference

#### Character Range

2

1.20%

#### Coverage

39774 - 40261

I don't think so. I think I've told you – I mean what I really wanted to say was that thing about democratization. I think that was great. I mean I am very impressed with that, and we're relatively new at

[A3's

institution] in astronomy. It's not an established place where astronomy has been done for 60 or 70 years. We

started only 30 years ago, and we would never have been in that pre print circuit. So from that point of view, it's

vital for starting a new research group in the area.

#### **Total References**

#### **Coverage**

2

#### **Total Users**

2.87%

1

#### **Internals\interviews\arXiv users\09020403\_a1**

#### **Document**

#### **References**

#### **Coverage**

1

#### **Tree Nodes\Themes\Democratization of the scholarly discourse\value for smaller institutions**

#### **Node Coding**

1.25%

#### Reference

#### Character Range

1

1.25%

#### Coverage

5862 -

6684

Well, it's completely changed the way communication occurs in the literature. I'm old enough to have piles of preprints as you can see some of my old piles. So it used to be that when you were at a big institution like Princeton or Harvard, you had an advantage because those places got all the recent results, all the preprints and all and so it's been much more – it's now much more egalitarian. Each – we as scientists anxiously – as we are finishing a work, we're anxiously thinking, "How quickly can I post this to astro-ph?" And what will people think? And what will people say? And often the reaction is less exciting as you hoped for, but nonetheless, it's out there. It's the sort of thing that you have no excuse – you're a scientist – not to know a current piece of work. You just don't have any excuse.

#### **References**

#### **Coverage**

1

#### **Tree Nodes\Themes\Impact on scholarly process\collaborative work**

#### **Node Coding**

0.46%

#### Reference

#### Character Range

1

0.46%

#### Coverage

31317 - 31617

So, I also find out –

I have a good network of colleagues and I go to conferences and so you learn about what's going on at conferences from colleagues and that will then feed me into – that will then –

I'll come back and I'll

download the arXiv if I've missed it or better yet, the published paper.

#### **References**

**Coverage**

2

**Tree Nodes\Themes\Integration with scholarly  
context\arXiv, ADS, part of everyday life  
Node Coding**

**2.00%**

*Reference*

*Character Range*

1

0.98%

Coverage

46777 -

47419

Well, I think every single experiment – every single project that I do is of that sort – that being able to sit at my desk and access all of these resources immediately without cost. Unfortunately, I don't appreciate how wonderful it is. But it's like this all the time. The reports of new targets are released. You can get – you may not have heard about this, but you can get alerts on your blackberry from the NASA satellite telling you of the currents of the Gamma Ray Burst.

And that's open access and you can sign up for that feature. So, I'm not amazed at it any more. It's a little sad, I mean I should be more amazed, but –

*Reference*

*Character Range*

2

1.02%

Coverage

47475 - 48146

It has evolved – because I grew – I came from the days of dedicated –

a dedicated typesetting word processing

unit where you typed on a specific unit, there was a printer attached to it and I also used the typewriter when I

was a kid in high school and long typing. And so I have watched as e-

mail became this thing that was unstable

and "Ooh, maybe it'll work today and maybe it won't" to something where, again, there's no excuse for saying,

"Oh, I didn't get that e-

mail." Because 99.99 percent of the time they go through. So, these tools are just a

natural part of what I do and I wouldn't be able to do as much as I do without them. So, I don't have the aha

–

**Total References**

**Coverage**

4

**Total Users**

**1.24%**

1

**Internals\interviews\PhilSci users\09022401\_ph26**

**Document**

**References**

**Coverage**

6

**Tree Nodes\Themes\Impact on scholarly  
output\access to things before you would have  
otherwise**

**Node Coding**

**2.86%**

*Reference*

*Character Range*

1

0.30%

Coverage

9119 -

9243

Well, I mean I don't know whether it's more often – I mean it's – you have access to things before you would have otherwise.

*Reference*

Character Range	
2	
0.37%	
Coverage	
9283 - 9435	
I mean I only have access there to papers that only appear two years down the line, and if it's a topic I'm really working on, that's a huge difference.	
Reference	
Character Range	
3	
0.52%	
Coverage	
9489 - 9702	
I mean two years is half a lifetime a graduate student, so it is very important to have things as soon as possible, so that is really a great help. I think the main advantage of these things is that they're fast.	
Reference	
Character Range	
4	
0.52%	
Coverage	
25789 - 26005	
No, it's not – not quantity –	
no, as I said before, I guess I'm repeating my point here. I mean the main use I see is that you just get things before they come out in journals. It is really being ahead of the time.	
Reference	
Character Range	
5	
0.30%	
Coverage	
26610 - 26735	
It is really – you look at them because you want to be sure you're not trying to do the same thing as someone has just done.	
Reference	
Character Range	
6	
0.85%	
Coverage	
26807 - 27159	
That is really the only use they have systematically, just to make sure you don't reinvent the wheel because as often that case that their topic is sort of in the air. I mean ideas usually have a context; they don't materialize out of nothing. And there's always the worry that what you're just trying to do has actually just been done by someone.	
Total References	
Coverage	6
Total Users	2.86%
	1
Internals\interviews\PhilSci users\09022402_ph32	
Document	
References	
Coverage	1
Tree Nodes\Themes\Democratization of the scholarly discourse\value for smaller institutions	
Node Coding	
1.01%	
Reference	
Character Range	
1	
1.01%	

Coverage

37997 -

38640

Yeah. I would say it's clearly helped me a great deal because at the small school I'm at, they just don't have many – they don't have many journals that they keep a paper copy there. And so I would have had to order inter-library loan and gotten paper copies, and that would have definitely slowed me up, so I would say that it's definitely improved the quality of the work I do and the ease of doing it. And so I definitely would say I've benefited from it. In terms of how easy they are to use online, I think the first time you – couple times you always use it, it's kind of slow because you don't quite know how to navigate through it.

#### **References**

##### **Coverage**

3

**Tree Nodes\Themes\Impact on scholarly  
output\access to things before you would have  
otherwise**

#### **Node Coding**

**1.87%**

*Reference*

*Character Range*

1

0.92%

Coverage

1339 - 1929

And also I would say the other one I use the most would be the PhilSci Archive because I do philosophy o  
f

science and a lot of the articles that will be coming out –

or earlier versions of articles that'll be coming out – are

put up on that site, and those are easy to obtain right over –

easily as well. And you can then kinda get an idea

because they have sort of an update page of things that people are putting up there. You can kinda get  
an idea

of what other people are working on, so it kind of gives you a feel as to what –

what's happening in the field as

it were that I work in.

*Reference*

*Character Range*

2

0.79%

Coverage

20890 -

21394

So I think it has really helped the industry a great deal in terms of the ease of doing research and getting  
material. And now, like I said, that even though I really haven't quite used it as much as I plan to, the fact that  
all this historical stuff in terms of original manuscripts being photocopied and put up on the website. And the  
ability, then, to do searches through them for particular words or whatever, is going to make that kind of a  
really nitty gritty historical research so much easier.

*Reference*

*Character Range*

3

0.16%

Coverage

36596 - 36700

By the fact that you can learn the thing today versus five years from now, right?

Interviewee: Right.

#### **References**

##### **Coverage**

1

**Tree Nodes\Themes\Impact on scholarly  
output\serendipity**

#### **Node Coding**

**0.59%**

*Reference*

*Character Range*

1

0.59%

Coverage

42213 -

42592

Yeah, it's very serendipitous. Yeah, you don't - I mean there's always the possibility that you may just look at something and say, "Wow, you know that's -" and that really changes your thinking about a problem. And if it wasn't for the online access, you might not have had that moment, so I mean it's clear - it clearly I think is - you know, made it a much better situation.

## **References**

### **Coverage**

5

**Tree Nodes\Themes\Impact on scholarly process\PhilSci, place to swap ideas**

### **Node Coding**

**2.71%**

*Reference*

*Character Range*

1

0.41%

Coverage

4498 - 4758

You know, it's really convenient in that regard, but PhilSci Archive I really see more as kind of like - it's kinda

like a place where we kinda swap our new ideas or we try them out. And so it has a little bit more of a development - developing aspect to it.

*Reference*

*Character Range*

2

0.83%

Coverage

6920 -

7453

Yeah, I think it's a - I think it's really, really a great thing. I think it's really helped the community to be aware of what we're all doing. I mean I guess that was part of the reason we had conferences for so long, was for that purpose that we could get together and share ideas and talk about the stuff we're into and try out things. And it's in many ways - open access is sort of like doing the same thing only much easier and at your own pace. And so I really think it's a - I think it's a really big and developing thing.

*Reference*

*Character Range*

3

0.88%

Coverage

54660 - 55224

Yeah, that I certainly -

that's been one of the nice things I think in that it's kept people more connected on what people are doing. So yeah, I would say that's been a great -

I mean that certainly changed and made things -

what would be the word for it? It's a new way to think about the profession. It gives you sort of a new

-

it's sort of a nice tool to know where information is being sort of exchanged, whereas before like I said, i

t

would have to be word of mouth, maybe at a conference. It's given sort of a site that you can go to to look for things.

*Reference*

*Character Range*

4

0.54%

Coverage

58318 -

58661

And I think actually put like a PhilSci link there, and I've seen PhilSci links in other journals - on published articles in journals where the link - so yeah, so I mean there is a sense in which even though I always see it less as a place - I don't really see it quite as a publishing place. I see it as more of like an information kind of -

*Reference*

*Character Range*

5  
0.05%  
Coverage  
58716 - 58746  
You know, collaboration place.

**References**  
**Coverage**

1

**Tree Nodes\Themes\Integration with scholarly  
context\PhilSci, not a publishing place**

**Node Coding**

**0.54%**  
*Reference*  
*Character Range*

1  
0.54%  
Coverage  
58318 -  
58661

And I think actually put like a PhilSci link there, and I've seen PhilSci links in other journals – on published articles in journals where the link – so yeah, so I mean there is a sense in which even though I always see it less as a place – I don't really see it quite as a publishing place. I see it as more of like an information kind of –

**Total References**  
**Coverage**

11

**Total Users**

1.34%

1

Coding Summary Report



## Appendix I: Summary of perceived values and implications for each theme

Table I1: Mapping and summary of the perceived values and implication of the properties for each theme for the arXiv group

Main themes / arXiv group	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>arXiv scientists</b>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- barriers of entry removed</li> <li>- new interactions possible</li> <li>- collaboration</li> <li>- trigger new ideas</li> <li>- serendipity</li> <li>- non peer-reviewed knowledge can enter the <b>knowledge network</b></li> <li>- enables article visibility, discoverability, access</li> <li>- scholar's visibility increased</li> <li>- accelerated research</li> <li>- latest research always immediately available</li> <li>- citation impact increase</li> <li>- speedup the search process for article</li> </ul>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- serendipity</li> <li>- more complete and well informed research</li> <li>- unpublished knowledge triggers new ideas</li> <li>- trigger new ideas</li> <li>- motivate research</li> <li>- better quality articles produced</li> </ul>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- new actors can emerge</li> <li>- commercial journals no longer visited</li> <li>- realigned knowledge production context</li> <li>- article visibility increased</li> <li>- "morning coffee" emerges as structure</li> <li>- critical venue for career boost</li> </ul>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- open the research process</li> <li>- open the knowledge ecosystem</li> </ul>

	<b>Time/early/quick</b> <ul style="list-style-type: none"> <li>- read early</li> <li>- read quick</li> <li>- distribute fast</li> <li>- knowledge from articles enters earlier</li> <li>- enable collaboration</li> <li>- rate of research enhanced</li> <li>- career jump-started</li> <li>- increased article visibility, discoverability, access</li> <li>- speedup the search process</li> </ul>	<b>Time/early/quick</b> <ul style="list-style-type: none"> <li>- new knowledge available early</li> <li>- trigger new ideas</li> <li>- realign <b>knowledge networks</b></li> <li>- relevant and cutting edge research possible</li> <li>- enhanced rate of research</li> <li>- unpublished knowledge triggers new ideas</li> <li>- new knowledge enters the ecosystem earlier</li> <li>- more time spent on research</li> <li>- able to analyze larger amounts of data</li> </ul>		
	<b>Space/wider distribution</b> <ul style="list-style-type: none"> <li>- knowledge distributed wide</li> <li>- location independent</li> <li>- enable collaboration</li> <li>- increased article visibility, discoverability, access</li> </ul>	<b>Space/wider distribution</b> <ul style="list-style-type: none"> <li>- serendipity</li> <li>- trigger new ideas</li> <li>- relevant and cutting edge research possible</li> </ul>		
	<b>Integration</b> <ul style="list-style-type: none"> <li>- reduces publication gap</li> <li>- makes arXiv, ADS, Chandra part of knowledge production</li> <li>- collaboration</li> <li>- trigger new ideas</li> <li>- serendipity</li> </ul>	<b>Integration</b> <ul style="list-style-type: none"> <li>- enables serendipity</li> <li>- trigger new ideas</li> <li>- enhanced rate of research</li> <li>- unpublished knowledge triggers new ideas</li> <li>- motivate research</li> <li>- availability of raw data makes more complete research</li> <li>- comprehensive view of a research problem</li> </ul>		

	<b>Linking</b> <ul style="list-style-type: none"> <li>- enhances article visibility, findability, accessibility</li> <li>- ADS visited more than arXiv</li> <li>- arXiv as channel for reading latest research</li> <li>- uniformity to the search process</li> <li>- accelerated research</li> </ul>	<b>Linked</b> <ul style="list-style-type: none"> <li>- enhances article visibility, findability, accessibility</li> <li>- availability of raw data makes more complete research</li> <li>- better quality articles produced</li> <li>- non peer-reviewed article enter the knowledge ecosystem</li> <li>- comprehensive view of a research problem</li> </ul>		
	<b>Inclusiveness</b> <ul style="list-style-type: none"> <li>- scanned older journal made accessible</li> <li>- trust to find latest research</li> <li>- trigger new ideas</li> <li>- collaboration</li> <li>- quality non-peer –review materials available</li> <li>- rate of research enhanced because of access</li> </ul>	<b>Inclusiveness</b> <ul style="list-style-type: none"> <li>- enables serendipity</li> <li>- scanned old journals become part of the <b>knowledge network</b></li> <li>- trigger new ideas</li> <li>- quality non-peer –review materials available</li> <li>- more complete and well informed research</li> <li>- motivate research</li> </ul>	<b>Inclusiveness</b> <ul style="list-style-type: none"> <li>- citation impact increases</li> </ul>	<b>Inclusiveness of unpublished knowledge</b> <ul style="list-style-type: none"> <li>- <b>knowledge network</b> realigned</li> </ul>
		<b>Access to latest research</b> <ul style="list-style-type: none"> <li>- new knowledge available early</li> <li>-</li> </ul>	<b>Early access to latest research</b> <ul style="list-style-type: none"> <li>- must be visited</li> </ul>	
		<b>Sub-categorization a challenge</b> <ul style="list-style-type: none"> <li>- inhibits serendipity</li> <li>- barrier to learning and discovery</li> <li>- extends the time to do research</li> </ul>		
			<b>Part everyday life</b> <ul style="list-style-type: none"> <li>- arXiv, ADS used</li> <li>- arXiv, ADS complement the peer review process</li> <li>- because it is trusted</li> </ul>	

		<p><b>Trust</b></p> <ul style="list-style-type: none"> <li>- ADS sufficient to find latest research</li> </ul>	<p><b>Trust</b></p> <ul style="list-style-type: none"> <li>- because of the pre-print culture</li> <li>- access to commercial journals marginalized</li> <li>- good enough proxies that define the disciplinary knowledge ecosystem</li> <li>- replaced the commercial journal</li> <li>- realigned knowledge production context</li> <li>- Critical venue for career boost</li> </ul>	
			<p><b>Central role</b></p> <ul style="list-style-type: none"> <li>- because of Speed, early access, inclusiveness, openness, time of distribution</li> <li>- no excuse not to know about the latest research</li> <li>- expected to read and post to arXiv</li> <li>- “morning coffee” emerges as structure</li> <li>- Critical venue for career boost</li> </ul>	
	-		<p><b>Completeness</b></p> <ul style="list-style-type: none"> <li>- replaced the commercial journal</li> <li>- everyday use</li> </ul>	

			<b>Total access</b> <ul style="list-style-type: none"> <li>- access to commercial journals marginalized</li> <li>- good enough proxies that define the disciplinary knowledge ecosystem</li> </ul>	
			<b>Related to pre-print culture</b> <ul style="list-style-type: none"> <li>- traditional pre-print culture transposed into new realm</li> </ul>	
				<b>Smaller institutions can participate</b> <ul style="list-style-type: none"> <li>- research process realigned</li> </ul>
				<b>Bypass peer-review</b> <ul style="list-style-type: none"> <li>- quality knowledge can be produced even by bypassing of the peer-review process</li> <li>- “liberation” from the peer-review bias</li> </ul>
				<b>Enable new entrants</b> <ul style="list-style-type: none"> <li>- smaller departments from well off institutions can also participate</li> <li>- new scholars can participate</li> </ul>

Table I2: Mapping and summary of the perceived values and implication of the properties for each theme for the PhilSci scholars

Main themes / PhilSci group	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
<b>PhilSci group of scholars</b>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- access to up-to-date research (some of which is unpublished )</li> <li>- displacing commercial journal's traditional role</li> <li>- share conference proceedings and pre-prints before a conference takes place</li> <li>- reduce publication gap</li> <li>- enables scholarly communication</li> <li>- enable the exchange of materials that are not formally published</li> <li>- PhilSci with the role of a conference, a collaborative tool, exchange ideas and thoughts,</li> <li>- Integration, Enables browsing mentality and serendipity</li> <li>- easier for researcher in finding, discovering, accessing articles</li> <li>- inclusion of older docs encourage new research</li> <li>- speeding up research</li> </ul>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- quicker, faster access to latest research</li> <li>- share conference proceedings and pre-prints before a conference takes place</li> <li>- source of new ideas</li> <li>- collaboration</li> <li>- preparatory place for articles that would be published in journals</li> <li>- faster distribution of own research</li> <li>- being ahead of time</li> <li>- enable the exchange of materials that are not formally published</li> <li>- a collaboratory</li> <li>- new and unpublished ideas may enter the knowledge ecosystem</li> <li>- better framing of research problems</li> <li>- serendipity</li> <li>- new type of research enabled/encouraged because knowledge enters the ecosystem that otherwise would not have entered</li> <li>- speeding up research</li> </ul>	<p><b>Openness</b></p> <ul style="list-style-type: none"> <li>- wider, quicker distribution of latest research materials</li> <li>- enables to see completeness of scholars' research</li> <li>- enables forward looking view of what's coming up in the discipline</li> <li>- yet not very crucial for knowledge production</li> <li>- find relevant articles</li> </ul>	<p><b>Smaller institutions can participate</b></p> <ul style="list-style-type: none"> <li>- emphasized by one scholar only</li> </ul>

Main themes / PhilSci group	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
	<b>Time/fast/quick</b> <ul style="list-style-type: none"> <li>- faster access to latest research</li> <li>- share conference proceedings and pre-prints before a conference takes place</li> <li>- enhanced scholarly publishing process</li> <li>- shorten publication lag</li> <li>- enables the scholars not to reinvent the wheel</li> <li>- enabling the authors to disseminate his work for other to read and influence their research</li> <li>- able to do more research</li> </ul>	<b>Time/quick/fast</b> <ul style="list-style-type: none"> <li>- quicker, faster access to latest research (not a necessary condition for writing articles)</li> <li>- preparatory place for articles that would be published in journals</li> <li>- faster distribution of own research</li> <li>- access to latest research before they are published by commercial journals</li> <li>- enables the scholars not to reinvent the wheel</li> <li>- up-to-date with latest research</li> <li>- better article produced because of ability to do more</li> <li>- better quality of articles produced</li> </ul>	<b>Time/quick/fast</b> <ul style="list-style-type: none"> <li>- wider, quicker distribution of latest research materials</li> <li>- enables to see completeness of scholars' research</li> </ul>	
	<b>Space/wide</b> <ul style="list-style-type: none"> <li>- wider distribution of the latest research</li> <li>- enhanced scholarly publishing process</li> <li>- enhances scholarly research process</li> <li>- enabling the authors to disseminate his work for other to read and influence their research</li> </ul>	<b>Space/wide</b> <ul style="list-style-type: none"> <li>- access to latest research before they are published by commercial journals</li> <li>- enhances scholarly research process</li> </ul>	<b>Space/wide</b> <ul style="list-style-type: none"> <li>- wider, quicker distribution of latest research materials</li> <li>- enables to see completeness of scholars' research</li> </ul>	

Main themes / PhilSci group	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
	<b>Some integration with conferences</b> <ul style="list-style-type: none"> <li>- a collaborative tool</li> <li>- exchange ideas for conferences</li> <li>- complementing the conference structure</li> </ul>	<b>Integration</b> <ul style="list-style-type: none"> <li>- collaboration</li> <li>- connect scholars with each other</li> <li>- a conference, collaboratory to exchange ideas and thoughts</li> <li>- new and unpublished knowledge may enter the scholarly ecosystem</li> </ul>		
	<b>Access to latest research</b> <ul style="list-style-type: none"> <li>- up-to-date research (some of which is unpublished )</li> <li>- reduce publication gap</li> <li>- enables the scholars not to reinvent the wheel</li> </ul>	<b>Early access to latest research</b> <ul style="list-style-type: none"> <li>- share conference proceedings and pre-prints before a conference takes place</li> <li>- source of new ideas</li> <li>- access to latest research, some might not be published</li> <li>- enables the scholars not to reinvent the wheel</li> <li>- being ahead of time</li> <li>- enables view of latest research</li> <li>- better quality of articles produced</li> <li>-</li> </ul>		
	<b>Collaboration / collaboratory</b> <ul style="list-style-type: none"> <li>- publishing venue for conference papers</li> <li>- connect scholars with each other</li> </ul>		<b>Collaboratory</b> <ul style="list-style-type: none"> <li>- transitory place for articles on their way to peer-review and publication</li> </ul>	
	<b>Home pages as access tools</b> <ul style="list-style-type: none"> <li>- Home Pages as critical actors in scholars' production process</li> <li>- Used to access pre-prints</li> <li>- Find the latest research</li> </ul>		<b>Home pages and other access tools</b> <ul style="list-style-type: none"> <li>- integrate with scholarly process</li> <li>- means for accessing works of others</li> <li>- as access tools</li> </ul>	



Main themes / PhilSci group	Impact on scholarly process	Impact on scholarly output	Integration with scholarly context	Democratization of the scholarly discourse
			- Google used as access tool	
		<b>Some role for invisible colleges</b> <ul style="list-style-type: none"> <li>- enables to get direct access to pre-prints before they are publishes</li> <li>- restrictive to serendipity</li> <li>-</li> </ul>	<b>Limited role</b> <ul style="list-style-type: none"> <li>- time the most important factor</li> <li>- and access to latest research</li> <li>- and wider dissemination</li> <li>- Isolated (“no life on its own”)</li> <li>- Limited value, information processing value only</li> <li>- Limited role even with the early availability of latest research (authors exchange via Home Pages)</li> <li>- integrated in the scholarly context of one particular scholar, not so across the discipline</li> </ul>	
			<b>No central role</b> <ul style="list-style-type: none"> <li>- perceived no direct value at personal or disciplinary level</li> <li>- PhilSci not central to scholars’ activities on daily basis</li> </ul>	
			<b>Pre-print culture</b> <ul style="list-style-type: none"> <li>- old paper based transposed into digital realm</li> </ul>	

## Bibliography

- About arXiv. Retrieved March 25, 2008, from <http://arxiv.org/help/general>
- Abramson, B. D. (1998). Translating nations: Actor-network theory in/and Canada. *Canadian Review of Sociology and Anthropology*, 35(1), 1-20.
- Akrich, M. & Latour, B. (1997). A summary of a convenient vocabulary for the semiotics of human and nonhuman assemblies. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 259-264). Cambridge, MA: The MIT Press.
- Antelman, K. (2004). Do Open-Access Articles Have a Greater Research Impact? *College & Research Libraries News* 65 (5), 372-382. Retrieved on October 6, 2005, from [http://eprints.rclis.org/archive/00002309/01/do\\_open\\_access\\_CRL.pdf](http://eprints.rclis.org/archive/00002309/01/do_open_access_CRL.pdf)
- Antelman, K. (2006). Self-archiving practice and the influence of publisher policies in the social sciences. *Learned Publishing* 19, (2) 85-95.
- arXiv. Retrieved March 25, 2008, from <http://www.arxiv.org>
- arXiv Primer. Retrieved March 25, 2008, from <http://arxiv.org/help/primer>
- Bailey, C.J. (2005). Open Access Bibliography. Washington, D.C: Association of Research Libraries. Retrieved on September 25 2006, from <http://www.escholarlypub.com/oab/oab.pdf>
- Baker, C., Wuest, J. & Stern, P.N. (1992). Method slurring: the grounded theory/phenomenology example. *Journal of Advanced Nursing* (17), 1355-1360.
- Barry, C.A. (1995). Critical Issues in Evaluating the Impact of IT on Information Activity in Academic Research: Developing a Qualitative Research Solution. *LISR* 17, 107-134.
- Bassi, L. (1997). Harnessing the Power of intellectual capital. *Training & Development*, 51(12), 25-28.
- Bates, M.J. (1994). The design of databases and other information resources for humanities scholars: The Getty online searching project report no.4. *Online and CDROM Review*, 18 (6), 331-340.
- Bijker, W. E. & Law, J. (1997). General introduction. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 1-19). Cambridge, MA: The MIT Press.

- BOAI. (2002). Budapest Open Access Initiative. Retrieved November 5, 2006 from <http://www.soros.org/openaccess/>
- Borgman, C. (1989). Bibliometrics and Scholarly Communication. *Communication Research* 16, 583-99.
- Borgman, C. L. (1999). What are digital libraries? Competing visions. *Information Processing & Management*, 35 (3), 227-243.
- Brody, T.D. (2006, May). Evaluating Research Impact through Open Access to Scholarly Communication. *Doctoral Dissertation, University of Southampton*. Retrieved December 28, 2007, from <http://eprints.ecs.soton.ac.uk/13313/1/brody.pdf>
- Buckland, M. (1991). *Information and Information Systems*. Chapters 1, 4, 5 & 6. New York: Praeger
- Brown, C. (2001). The E-volution of Preprints in the Scholarly Communication of Physicists and Astronomers. *Journal of the American Society for Information Science and Technology* 52 (3), 187-200.
- Chan, L. (2004). Supporting and Enhancing Scholarship in the Digital Age: The Role of Open-Access Institutional Repositories. *Canadian Journal of Communication* 29, (3). Retrieved January 1, 2007 from <http://www.cjc-online.ca/viewarticle.php?id=850>
- Citeseer: Accessed November 12, 2006, at <http://citeseer.ist.psu.edu/>
- Coleman, A., Bracke, P. & Karthik, S. (2004). Integration of Non-OAI Resources for Federated Searching in DLIST, an Eprints Repository. *D-Lib Magazine* 10 (7/8). Retrieved November 12, 2006, from <http://www.dlib.org/dlib/july04/coleman/07coleman.html>
- Coleman, A. & Roback, J. (2005). Open Access Federation for Library and Information Science: dLIST and DL-Harvest. *D-Lib Magazine*. Retrieved November 5, 2006 from <http://www.dlib.org/dlib/december05/coleman/12coleman.html>
- Covi, L. M. and Kling, R. (1996). Organizational dimensions of effective digital library use: Closed rational and open natural systems models. *Journal of the American Society for Information Science* 47 (9), 672-689.
- Creswell, J.W. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Thousand Oaks, CA: Sage Publications, Inc.
- Creswell, J.W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage Publications, Inc.

- Davis, P.M. & Connolly, M.J.L. (2007). Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University's Installation of DSpace. *D-Lib Magazine*, 13 (3/4). Retrieved October 10, 2009, from <http://www.dlib.org/dlib/march07/davis/03davis.html>
- DeSanctis, G. & Poole, M. S. (1994). Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science*, 5 (2), 121-147
- dLIST. Digital Library of Information Science and Technology. Retrieved November 12, 2006, from <http://dlist.sir.arizona.edu/>
- DSpace. Accessed March 25, 2008, at <http://www.dspace.org/>
- Ducheneaut, N. (2005). Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Computer Supported Cooperative Work* (14), 323–368.
- E-LIS. The open archive for Library and Information Science. Accessed November 12, 2006, at <http://eprints.rclis.org/>
- EPrints. Accessed March 25, 2008, at <http://www.eprints.org/>
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin* 54 (4), 327-358.
- Fry, J. (2006). Scholarly research and information practices: a domain analytic approach. *Information Processing & Management* 42, 299-316.
- Google Scholar. Accessed November 12, 2006, at <http://scholar.google.com/>
- Greestone. Accessed March 25, 2008, at <http://www.greenstone.org/>
- Harnad, S. (2006) Maximizing Research Impact Through Institutional and National Open-Access Self-Archiving Mandates. In K. Jeffrey (Eds.) *Proceedings of CRIS2006. Current Research Information Systems: Open Access Institutional Repositories (in press)*, Bergen, Norway. Retrieved September 9, 2006 from <http://eprints.ecs.soton.ac.uk/12093/02/harnad-crisrev.pdf>
- Hjørland, B., & Albrechtsen, A. (1995). Toward a New Horizon in Information Science: Domain-Analysis. *Journal of the American Society for Information Science* 46 (6), 400–425.
- Hedlund, T., Gustafsson, T., and Björk, B. (2004). The Open Access Scientific Journal: An Empirical Study. *Learned Publishing* 17 (3), 99–209. Retrieved September 12, 2006 from <http://oacs.shh.fi/publications/199-210.pdf>

- Hunter, D. (2005). Walled Gardens. *Washington & Lee Law Review*, 62, 1-28.  
(Using a pre-print version Public beta Revision 1.1, Tuesday, December 14, 2004).
- JSTOR. Accessed on April 14, 2010, at <http://www.jstor.org>
- Kennan, M.A. (2008). Reassembling Scholarly Publishing: Open Access, Institutional Repositories and the Process of Change (Unpublished doctoral dissertation). The University of New South Wales. Retrieved September 7, 2009 from <http://ils.unc.edu/~wildem/ASIST2009/Kennan.pdf>
- Kling, R. (2003). The Internet and Unrefereed Scholarly Publishing. In B. Cronin (Eds.), *Annual Review of Information Science and Technology* (38) (pp. 591–631). Medford, NJ: Information Today, Inc.
- Kling, R., & Callahan, E. (2003). Electronic journals, the internet, and scholarly communication. In B. Cronin & D. Shaw (Eds.), *Annual review of information science and technology* 37, 127-177. Medford, NJ: Information Today, Inc.
- Kling, R. & Covi, L. M. (1995). Electronic Journals and Legitimate Media in the Systems of Scholarly Communication. *The Information Society* 11 (4), 261-271. Retrieved September 12, 2006 from <http://www.slis.indiana.edu/TIS/articles/klingej2.html>
- Kling, R. & McKim, G. (1999). Scholarly communication and the continuum of electronic publishing. *Journal of the American Society for Information Science* 50 (10), 890-906.
- Kling, R., and McKim, G. (2000). Not Just a Matter of Time: Field Differences and the Shaping of Electronic Media in Supporting Scientific Communication. *Journal of the American Society for Information Science* 51 (14), 1306–1320.
- Kling, R., McKim, G. and King, A. (2003). A Bit More to It: Scholarly Communication Forums as Socio-Technical Interaction Networks. *Journal of the American Society for Information Science and Technology* 54 (1), 47– 67.
- Knorr Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press
- Latour, B. (1999). On recalling ANT. In J. Law & J. Hassard (Eds.), *Actor network theory and after* (pp. 14-25). Malden, MA: Blackwell.
- Latour, B. (2005). *Reassembling the Social: An introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Law, J. (1999). After ANT: Complexity, naming and topology. In J. Law & J. Hassard (Eds.), *Actor network theory and after* (pp. 1-14). Malden, MA: Blackwell.

- Law, J. & Bijker, W. E. (1997). Postscript: Technology, Stability, and Social Theory. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 290-308). Cambridge, MA: The MIT Press.
- Law, J. & Callon, M. (1997). The Life and Death of an Aircraft: A Network Analysis of Technical Change. In W. E. Bijker & J. Law (Eds.), *Shaping technology/building society: Studies in sociotechnical change* (pp. 21-52). Cambridge, MA: The MIT Press.
- Law, J. & Hassard, J. (Eds.). (1999). *Actor network theory and after*. Malden, MA: Blackwell.
- Lawal, I. (2002). Scholarly Communication: The Use and Non-Use of E-Print Archives for the Dissemination of Scientific Information. *Issues in Science and Technology Librarianship*, 36. Retrieved September 14, 2005, from <http://www.istl.org/02-fall/article3.html>
- Littlejohn, S. (1992). *Theories of Human Communication*. Belmont, CA: Wadsworth
- McDonald, J.D. (2007). Understanding Journal Usage: A Statistical Analysis of Citation and Use. *Journal of the American Society for Information Science and Technology* 58 (1), 39–50.
- Miettinen, R. (1999). The Riddle of Things: Activity Theory and Actor-Network Theory as Approaches to Studying Innovation. *Mind, Culture and Activity*, 6(3), 170-195.
- Meyrowitz, J. (1994). Medium theory. In D. Crowley and D. Mitchell (Eds.). *Communication theory today* (pp. 50-77). Stanford, CA: Stanford University Press.
- Miles, M.B. & Huberman, A.M. (1994). *An Expanded Sourcebook: Qualitative Data Analysis*. Thousand Oaks, CA: Sage Publications, Inc.
- Mueller, J. (2008, April 23). Retiring support for OAI-PMH in Sitemaps. Retrieved September 24, 2009, from <http://googlewebmastercentral.blogspot.com/2008/04/retiring-support-for-oai-pmh-in.html>
- OAI. Retrieved November 11, 2006, from <http://www.openarchives.org/>
- OAI. *Tools*. Retrieved November 11, 2006, from <http://www.openarchives.org/tools/tools.html>
- OAI: F.A.Q. Retrieved November 11, 2006, from <http://www.openarchives.org/documents/FAQ.html>

- OAster. Accessed April 14, 2010, at <http://oaister.worldcat.org/>
- Okerson, A. (2000). Are we there yet? Online E-Resources Ten years After. *Library Trends* 48(4), 671-693.
- Palmer, C.L. (2005). Scholarly Work and the Shaping of Digital Access. *Journal of the American Society for Information Science and Technology* 56(11), 1140-1153.
- Palmer, C.L. & Cragin, M.H. (2008). Scholarship and Disciplinary Practices. In B. Cronin (Eds.), *Annual Review of Information Science and Technology* (42) (pp. 165–212). Medford, NJ: Information Today, Inc.
- PhilSci. Retrieved March 26, 2008, from <http://philsci-archive.pitt.edu>
- PhilSci Policy. Retrieved March 26, 2008, <http://philsci-archive.pitt.edu/policy.html>
- Ruhleder, K. (1994). Rich and Lean Representations of Information for Knowledge Work: The Role of Computing Packages in the Work of Classical Scholars. *ACM Transactions on Information Systems* (12) 2, 208-230.
- Savolainen, R. (2007). Information Behavior and Information Practice: Reviewing the “Umbrella Concepts” of Information-Seeking Studies. *Library Quarterly* 77 (2), 109-132.
- Selden, L. (2005). On Grounded Theory - With Some Malice. *Journal of Documentation* 61 (1), 114-129.
- Strauss, A. & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA: Sage Publications, Inc.
- Suleman, H. & Fox, E. (2002). The Open Archives Initiative: Realizing Simple and Effective Digital Library Interoperability. *Journal of Library Administration* 35 (½), 125–45.
- Talja, S., Keso, H., & Pietilainen, T. (1999). The production of `context' in information seeking research: a metatheoretical view. *Information Processing and Management* 35, 751-763
- Talja, S., Tuominen, K., & Savolainen, R. (2005). “Isms” in information science: constructivism, collectivism, and constructionism. *Journals of Documentation* 61 (1), 79-101.
- Van de Sompel, H., Payette, S., Ericksson, J., Lagoze, C., & Warner, S. (2004). Rethinking Scholarly Communication: Building the System that Scholars Deserve.

- D-Lib Magazine* 10 (9). Retrieved August 29, 2006 from <http://www.dlib.org/dlib/september04/vandesompe/09vandesompe.html>
- Van House, N. A. (2003). Digital libraries and collaborative knowledge construction. In A P. Bishop, B. P. Battenfield & N. A. Van House (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 271-295). Cambridge, MA: The MIT Press
- Warr, W.A. (2003). Evaluation of an Experimental Chemistry Preprint Server. *Journal of Chemical Information and Computer Sciences* 43 (2), 362–373.
- Whitley (2000). *The intellectual and social organization of the sciences* (2nd ed.). Oxford: Clarendon Press.
- Wick, C. (2001). Measuring Intellectual Capital at EDS. *Knowledge Management Review*, 4(5), 14-17.
- Willinsky, J. (2005). The unacknowledged convergence of open source, open access, and open science. *First Monday*, 10(8). Retrieved September 21, 2005, from [http://firstmonday.org/issues/issue10\\_8/willinsky/](http://firstmonday.org/issues/issue10_8/willinsky/)
- Willinsky, J. (2006). *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: The MIT Press.
- Wilson, T.D. (2006). Information, communication and society. *Svensk Biblioteksforskning* 15 (2). Retrieved November 4, 2006 from <http://www.hb.se/bhs/svbf/arkiv/wilson.pdf>
- Wimpenny, P. & Gass, J. (2000). Interviewing in phenomenology and grounded theory: is there a difference? *Journal of Advanced Nursing* 31 (6), 1485-1492.
- Wulff, J.L. & Nixon, N.D. (2004). Quality markers and use of electronic journals in an academic health sciences library. *J Med Libr Assoc* 92 (3), 315-322.
- Xia, J. & Sun, L. (2007). Assessment of Self-Archiving in Institutional Repositories: Depositorship and Full-Text Availability. *Serials Review* 33, 14-21.
- Yue, W., and Wilson, C.S. (2004). Measuring the citation impact of research journals in clinical neurology: A structural equation modeling analysis. *Scientometrics* 60 (3), 317-332.
- Zhang, Y. (2001). Scholarly Use of Internet-Based Electronic Resources. *Journal of the American Society for Information Science and Technology* 52 (8), 628–654.



## Curriculum Vita

### Mentor Cana

#### Education

- 2010      Ph.D., Information Science  
Graduate Program in Communication, Information and Library studies  
Rutgers, The State University of New Jersey
- 1997      Master Degree, Electrical Engineering / Telecommunication  
Stevens Institute of Technology, Hoboken, NJ
- 1996      Bachelors of Science, Electrical Engineering  
SUNY New Paltz, New Paltz, NY

#### Professional Experience

- 2005 –      Sr. Technology Solutions Professional, Media and Entertainment  
current      Microsoft Corp.
- 2005      Business Systems Analyst / Systems Architect / Consultant  
SBC / IBM
- 2004 -      Systems Engineer / Consultant, AT&T Labs  
2005      AT&T / Tekmark
- 2003 -      Teaching Assistant / Ph.D. Student  
2004      RUTGERS UNIVERSITY  
SCILS – The School of Communication, Information, and Library  
Studies
- 2002 -      Systems Engineer / Consultant  
2003      TELCORDIA / ATS
- 2000 -      Systems Analyst / Consultant, AT&T Labs  
2002      AT&T / LTI
- 2000      Senior Sales Systems Engineer, Internet Telephony  
ITXC Corp.
- 1998 -      Systems Architect / Consultant, AT&T Enterprise Architecture  
2000      AT&T

1997 - Engineer, Professional Services  
 1998 Bellcore / Telcordia

1995 - Teaching Assistant / Graduate Student  
 1997 Stevens Institute of Technology

#### Publications

Cana, M. (2009, May). *Scholars' interaction with self-archived open access (OA) repositories: Describing the performative agencies of actors in the socio-technological knowledge production topologies*. Poster submitted for presentation at Libraries in Digital Age (LIDA) 2009 Conference, Dubrovnik and Zadar, Croatia, May 25-30, 2009. (*Refereed International Conference*)

Dalbello, M., Saracevic, T., Palmer, C., and Cana, M. (2005, Sep 22). *Diffusion of knowledge in the field of digital library development: How is the field shaped by visionaries, engineers, and pragmatists?* Proceedings of the American Society for Information Science and Technology, USA, 564-566