SPECIFICATIONS
WMS Submission Tracking and Reporting

Since 2007, the Libraries have accepted faculty deposits into the institutional repository, RUcore, via its WMS (Workflow Management System). Between 2007 and 2013, a total of 267 submissions were processed into RUcore (38 per year on average, or roughly 3 per month), with about half of these coming from the ranks of the librarians. Throughout this period, there was no automated mechanism for tracking deposits, although the WMS administrators received email notifications when a deposit had been made, and later when it was ingested into the repository. The repository manager was largely reliant on these email notifications to track deposits, using one email folder for new deposit notifications, and moving each notification to a "Work completed" email folder when processing was done. When assistants were available to help process deposits, work was assigned by the manager, who continued to monitor when work was completed, using this email filing system.

In late 2014, the Open Access Policy implementation ramped up and processed deposits rose dramatically that year, from 49 in the previous year to 116. SOAR launched in January, 2015, with new functionalities and services that increased processing complexity; nonetheless processed deposits rose to 311 that year. We have 396 deposits thus far this year (through October 14, 2016), with 500 projected by calendar year end. The rate of return depositors has also increased dramatically after the launch of SOAR, and this year we have many instances of single authors depositing 10-15 works at once, which has implications for deposit processing.

In spring 2015, the SOAR Librarian was appointed, marking the first time more than one person was processing scholarship deposits into SOAR. She initiated a new system for assigning, tracking, status checking and reporting of deposits, using Excel spreadsheets. The spreadsheets also include a "tickle file" for reminders to check back with a depositor or publisher regarding a dormant request, or reminders to check for actual publication of an Accepted Manuscript so that we can add the "link to published version" that is so important to authors. The system provides far greater functionality than the email filing system, but still remains largely manual and somewhat cumbersome. As deposit rate goes up, the time spent maintaining the spreadsheet is rising accordingly, along with the possibility that additional assistance will be needed. When three people began simultaneously processing deposits this year, additional processing management issues arose. The repository manager was authorized to "claim" deposits, while the others with less experience were assigned deposits by her. Less experienced users are not authorized for all processes, so deposits would pass back and forth between processors more than once. It will continue to be the case that timely processing and effective management of the SOAR deposits will require more than one person working on multiple deposits simultaneously and a new system is needed to ensure processing is properly managed.

Regardless of who is doing the processing, the workflow, straightforward as it is, can start and stop at several points in the process. If the author deposits the wrong version, we request the correct version from the author and/or the publisher. Often a deposit is paused in midstream so that the processor can attend to another unrelated priority. On occasion, we will encounter an issue that requires intervention from IIS, such as a setting change on the PDF server or to enable processing of a new file format.

Currently, the starting point for tracking deposits is a pair of emails, delivered to a group email address, one announcing metadata has been submitted, the second announcing the accompanying file has been uploaded. Each processor on the group email list files these emails in whatever way s/he finds most effective for tracking. When the file is ingested, an "ingest complete" email is sent to the author, and copied to the group list. Processors with a high volume of email must be relied upon to properly flag deposit notifications so they aren't missed. To further complicate matters, occasionally-used alternative workflows do not trigger email notifications, for example, when an article's supplementary data is deposited before the article.

There is also no good way for one processor to know the status of a deposit that has been claimed by another processor. Although this information is recorded on the spreadsheet, sharing of that file carries an inherent risk

of inadvertent editing or loss. Our best option has been to record the status of a particular deposit in the metadata itself where it can be most easily seen by other processors, i.e. in the title field (for example, "Awaiting Accepted Manuscript from author"). This is cumbersome and has also prevented us from easily displaying the status of individual "in process" titles to users in their "My Account" pages, because title information is pulled from metadata, now doing double duty as a tracking mechanism.

With current deposit rates, it is no longer efficient or effective to rely on email notifications, individuals' idiosyncratic and private filing systems, and backend workarounds to quickly assess what work needs to be done, who's doing it, and how it should be prioritized. The time required to manually manage all deposits is becoming prohibitively high. An ideal solution would allow a processor to sit at her desk, open a WMS application, and see exactly what needs to be done, without having to sift through an email inbox, with all the distractions that can go along with that. In addition, we would like the ability to generate reports for management of deposits and for assessment of SOAR services.

This proposal is to develop an administrative portal to allow authorized users to claim, assign, track, monitor and report all deposits within a specified collection, in this case, the Rutgers University Faculty Collection that comprises SOAR. The feature could be extended to other collections now or later. The WMS Submission Tracking and Reporting feature would be designed to display new receipts and item status for all deposits upon login, with easy access from a dashboard to additional functions as listed below. The system should ensure historical data is retained through WMS cleanup.

CLAIMING & ASSIGNING FUNCTIONS

Allow authorized users to:

1. claim an item for themselves

2. assign an item to other authorized WMS users (this feature would not prevent others from working on the item; i.e. it is not an authorization system)

3. change a claim or reassign an item

4. view a list of all items, unclaimed, claimed, and assigned

5. filter the list of items by a particular claimant or assignee ("owner")

6. quickly access any resource listed under a particular claimant or assignee via a "quick link"

STATUS INDICATORS

1. Designate and display a "status" value for all items that have entered the system, from time of arrival through processing, ingest, and follow-up. There will be several status values, five or more; they should be configurable at any time, and include the following. Bracketed numbers are for ease of reference within this document and should not display.

   S1 Received ("new")
   S2 Pending author response
   S3 Pending publisher response
   S4 Editing in process
   S5 Under review
   S6 Ingested
   S7 Awaiting citation
   S8 Closed
   S9 Withdrawn
   SB Backlogged (a calculation)
   SP Pending author or publisher response (a calculation)

2. The user can change the Status value in either direction **except do not allow Ingest status to revert back to a previous status.** Timestamp the start and end point of each status, and calculate time in each status

3. Retain 'start' and 'end' timestamps associated with each status, for use in reports

4. Add a free text note field that captures processing notes and is only visible to WMS users. This information is not meant for descriptive metadata and only relates to information about the particular item's processing through the deposit/approval process. For example, the note might include information as to why an item is continuing in a particular status.

5. Allow the user to click through to the Existing Records list in Metadata and Digital Object when s/he clicks on the identifier in the dashboard. Clicking a WMS number on the dashboard listing will take the user to the WMS Record list for that collection (in Metadata and Digital Object) and then click the radio button for that record. Clicking on a PID in that same column will take the user to the corresponding record in RUcore.

6. Display the status value and owner next to the record in WMS's record list within the Metadata/Digital Object module

7. Allow easy access via a "quick link" to the item on the dashboard from the status value displayed in the Metadata/Digital Object module.

NOTIFICATIONS AND REMINDERS

1. Continue all email notifications (to WMS administrators) of SOAR deposit (both metadata submission and file upload). Continue notifications to depositors with cc to WMS administrators (Ingest, Most downloaded, Embargo has been lifted).

2. Set up timed notifications to authorized WMS users for the following scenarios:

   a. Item has entered the backlog (Criteria: S1 + S4 + S5 = >48 hours and neither S6 nor S9 were ever reached)

   > SUBJECT: SOAR resource has entered the backlog
   > BODY: The following item has entered the backlog: [citation; see Versioning specs]

   b. Author response has been pending more than two weeks (Criteria: time in S2 exceeds 2 weeks)

   > SUBJECT: SOAR resource processing is awaiting response from the author
   > BODY: Processing of the following item has been awaiting response from the author for more than two weeks: [citation; see Versioning specs]

   c. Publisher response has been pending more than two weeks (Criteria: time in S3 exceeds 2 weeks)

   > SUBJECT: SOAR resource processing is awaiting response from the publisher
   > BODY: Processing of the following item has been awaiting response from the publisher for more than two weeks: [citation; see Versioning specs]

   d. Item with incomplete citation has likely been published, so follow up (a complete citation and determination of publication date) is required to enable "link to published version" and allow the correct embargo to be set (Criteria: time in S7 exceeds 3 months. This notification should continue to be resent every 3 months the item remains in S7, until the status is changed to S8, Closed.)

   > SUBJECT: SOAR resource may be published; check for full citation and revise metadata
   > BODY: The following item has an incomplete citation; three months have passed since ingest and it may now be published: [citation; see Versioning specs]

FILTERING AND EXPORTING

1. Enable filtering and exporting to allow searching and reporting in support of workflow management and assessment, using data for all items and all sub-collections under a parent collection, and across collections. Examples include:

    a. number of deposits

    b. daily deposit rate

    c. ratio of ingests to submissions

    d. turnaround times

    e. number of items in each state, both within specified collections or subcollections and across collections

    f. average length of time any one item, or specified groups of items, were in each state during processing, to identify bottlenecks in the process

    g. number of items associated with claimants and assignees

    h. for a specified time period, snapshot indicating the number of items in each (or any) status (received, ingested, completed, withdrawn, backlogged, etc.)

2. Enable CSV export of data that would allow administrators to create their own reports, for example, in Excel.

Identifier	Title		Depositor	Status ▼	Owner ▼	Notes

Deposited [YYYY-MM-DD] to [YYYY-MM-DD]   Ingested [YYYY-MM-DD] to [YYYY-MM-DD]   **FILTER**

☐ Include turnaround time   ☐ Include time in [Status ▼]   **EXPORT**

Number of results: 3 of [total submitted to system]

Identifier ↕	Deposit Date ↕	Ingest Date ↕	Title ↕	Depositor ↕	Status ↕	Owner ↕	Notes ↕	Publisher ↕	Journal ↕	
81194	2017-01-10	2017-01-11	Best paper ever	Yang Yu	New	Jane Otto				edit
81193	2017-01-09		Okay paper	John Doe	Editing in process	Jane Otto	Files corrupted	T&F	CCQ	edit
81192	2017-01-08		Worst paper	Jane Doe	Pending author	Jane Otto	Publisher allows VoR			edit

Details about the "dashboard" display

"Number of results" indicates total results filtered of total resources received in the system.

If no dates are specified, include all resources that have entered the system.

Greyed text in boxes indicates a placeholder.

Black text in boxes indicates a pulldown.

Pulldowns are multiselects (i.e. one or more values can be selected within a pulldown) and you must be able to choose multiple filters. All filters are ANDed.

If time permits, for status and owner pulldowns, indicate in parens after the value the number of items with that value.

Column headings can be sorted in ascending or descending order by clicking the arrows. The status sort should be in sequential order by process, with the two calculated statuses at the end, i.e. ascending order would be New – Pending author – Pending publisher – Editing in process – Under review – Ingested – Awaiting citation – Closed – Withdrawn – Backlogged - Pending.

Allow the user to hide/unhide any column.

Clicking the 'edit' link opens a window (or new tab?) to allow editing of all values in that entry **except Identifier, Deposit Date, Title, and Depositor**. For 'Status' and 'Owner' values, you might display editable values buttons; hold down the button to get a pulldown and change the value there.

Export formatting

Clicking the Export button will export to a spreadsheet the results of the search. Include for each listed item a column for each of the following data points; definitions follow.

If time permits, allow the user to select data points for export. For example, the user might wish to see only items for a particular collection with ID, title, publisher, and journal.

> Identifier
>
> Title
>
> Deposit date
>
> Ingest date
>
> Depositor [collection]
>
> Depositor collection ID
>
> Original owner
>
> Other owner1, Other owner2, etc. (one column for each other owner not previously listed)

> | |
> |---|
> | Status1 |
> | Status1 start timestamp |
> | Status1 end timestamp |
> | Status1 time in Status1 |

> *[Repeat Status block of 4 values for each status, in sequential order, followed by Backlogged and Pending statuses, as applicable]*
>
> Notes
>
> Publisher
>
> Journal

Terminology

Deposit date Date the item entered the system, displayed as YYYY-MM-DD.

Depositor Name of the depositing author [collection], pulled from digiprov:

```
<rulib:digiprovEvent ID="DPEVENT-1">
        <rulib:type>Create metadata</rulib:type>
        <rulib:dateTime>2017-01-06</rulib:dateTime>
        <rulib:associatedEntity ID="DPAE-1">
                <rulib:name>Anderson Katherine</rulib:name>
                <rulib:role>Depositing Author</rulib:role>
                <rulib:affiliation>10173</rulib:affiliation>
                <rulib:email>katie.anderson@camden.rutgers.edu</rulib:email>
        <rulib:identifier>katander@rutgers.edu</rulib:identifier>
```

The depositor text box should accept depositor (collection) name or collection ID (in the event two authors have the same or similar names).

Depositor collection ID Collection ID for the depositing author.

Identifier WMS # or PID (rutgers-lib:) number as applicable. WMS # reverts to PID on ingest. Clicking a WMS number on the dashboard listing will take you to the WMS record list for that collection (in Metadata and Digital Object) and will click the radio button for that record. Clicking on a PID in that same column will take you to the corresponding record in RUcore

Ingest date Date the resource was ingested into Fedora, displayed as YYYY-MM-DD.

Journal Title of journal to which the paper was submitted or in which it was published. Keyed by librarian as needed; if not needed the field remains empty.

Notes Keyed by librarian as needed; if not needed the field remains empty. In filtering and exporting, this input field will generate a keyword search of the (administrative) notes fields created in this module.

Original owner The first claimant. The name displays as WMS user Firstname Lastname, e.g., Jane Otto.

Other owner(s) Any owner after the first, i.e. any subsequent assignee. In exporting, make a separate column for each owner, in sequence. The name displays as WMS user Firstname Lastname, e.g., Jane Otto.

Owner The current claimant or assignee responsible for processing at the time of the query. The cell is empty until the item is claimed. The name displays as WMS user Firstname Lastname, e.g., Jane Otto.

Publisher Publisher of the journal in which the paper was published. Keyed by librarian as needed; if not needed the field remains empty.

Status Current processing status. Two status values are assigned by the system: The default value when an item enters the system is NEW. One other value is assigned by the system: INGESTED. (On ingest, the WMS number flips to the PID across the system.) All other values are manually changed by the librarian as each phase of work is complete. Consider the option of displaying Status value as a colored button to easily see at a glance what needs attention.
Status should be one of the following values; bracketed numbers are for ease of reference within this document and should not display. S1-S9 are sequential (although any status after the first can be bypassed); SB and SP are calculated.

   [S1] New - This is the default for items entering the system which have not yet been claimed.

   [S2] Pending author – Librarian has contacted author and is awaiting reply

   [S3] Pending publisher – Librarian has contacted publisher and is awaiting reply

[S4] Editing in process – Librarian is in the process of checking files, collection structure, and cover sheet, editing metadata, completing admin document processing, and affiliating co-author collections

[S5] Under review – One owner has completed work and has asked a supervisor to review it.

[S6] Ingested – The item has been ingested into Fedora

[S7] Awaiting citation – The ingested item lacks full citation metadata, which will be filled in after the item is published, with embargo adjusted as necessary.

[S8] Closed – Processing is complete; the ingested item has full citation metadata; the author has been informed of any lengthy embargo as applicable, and no more work is required.

[S9] Withdrawn – The librarian has deleted the item without fully processing it because it is a dupe, the author has requested it be withdrawn, it is not eligible for SOAR deposit, etc.

[SB] Backlogged – All deposits where S1 + S4 + S5 = 48h and neither S6 (ingest) nor S9 (withdrawn) was achieved.

[SP] Pending – All deposits with current status of S2 OR S3

Status#	In exporting, include a column for each status the item has reached, e.g., S1, S4, S6, S7, S8, SB. For each status reached, include a column for the Status value, Status start timestamp, Status end timestamp, and time in status. Use the name of the status, not the alphanumeric.
Status# start timestamp	The time the item enters that status.
Status# end timestamp	The time the item exits that status.
Status# time in status	End timestamp in that status minus Start timestamp in that status, expressed in minutes. If >120 minutes, express in hours; if >48 hours, express in days.
Time in status	End timestamp in the specified status minus Start timestamp in that status, expressed in minutes. If >120 minutes, express in hours; if >48 hours, express in days.
Title	Title of the item, pulled from descriptive metadata (mods:title).
Turnaround time	Time elapsed from time of deposit to time of ingest, expressed in minutes. If >120 minutes, express in hours; if >48 hours, express in days.

ADMINISTRATIVE FUNCTIONS

1. Allow addition of manual entries, including "legacy" entries from the SOAR Librarian's spreadsheet dating from February 18, 2014. As of January 10, 2017, this includes 946 entries. These could be added manually or imported from a spreadsheet if that functionality is available.

2. Retain a record of each item that has entered WMS, and the life cycle of that item's processing, for reference in WMS, even after that item has been purged from WMS.

3. If specified by the user in My Workflows, the module should open by default to the dashboard as shown above.

4. To integrate with WMS:

    a. Add a fifth heading/tab under the "Digital Object Workflow Management System" banner for "Tracking and Reporting," preceding the "Metadata and Digital Object" heading tab, so that people can move seamlessly between the Dashboard and the other modules (M&DO, Administration,

Configuration, User Account) as needed. This heading tab should only appear if the module is available for that collection.

b. In the Metadata and Digital Object module, on the "Existing Records" screen on the "Metadata and Digital Objects" tab:

    1. Change the "Status" column heading to "Ready for ingest"

    2. Add a column to the right of that for "Processing status"; this is where the status values will appear (New, Pending author, Pending publisher, etc.)

    3. Add a column to the right of "Processing status" for "Owner"

c. In WMS, under "My workflow," add an option to opt out of specific email notifications. Provide a list of notifications the administrator can check, as follows, and an opt-out button.

Notifications to administrators

☐ Metadata submission

☐ File upload

☐ Item has entered backlog

☐ Author has not responded in two weeks

☐ Publisher has not responded in two weeks

☐ Check for complete citation

Notifications to depositors, copied to administrators

☐ Ingest

☐ Most Downloaded

☐ Embargo has been lifted

Opt Out