Restricted Access to RUcore Collections and Resources: Requirements and Scenarios

Reviewed by Software-Architecture WG (November 6, 2014), and revised
Reviewed by User Services & Applications WG (November 13, 2014), and revised
Reviewed by CISC (November 26, 2014), and revised

Introduction

This document describes proposed approaches to restrict access to RUcore collections and resources. Currently, we are able to temporarily exclude access to a RUcore resource through the use of the XACML Policy that places an embargo on a datastream (e.g., a text document). This requirements document describes additional measures to restrict access to individual RUcore resources and collections. The measures are drawn from copyright holder requests, access requirement to comply with copyright law, and provisions of contracts including IRB decisions.

These are technical options for restricting access. The use of any of these options for an individual resource, collection, or groups of collections, will require a separate use and restrictions policy. For example, if the University (the Libraries together with the affected graduate school Dean's Offices) decides to apply one of the restriction options to ETDs, we will need a policy on whether that will be applied across the board or to an individual ETD upon request. In any case, we will clearly and unambiguously describe our restriction options to our users – copyright holders, librarians and staff, teaching and research faculty, and university administrators – in a separate statement.

Proposed Restricted Access Options

1. **Restrict to IP address or range**. A collection or single resource can only be accessed from a single IP address (such as a single computer workstation); or an IP address range such as a range of IP addresses in a single library building, or the range(s) of IP addresses that constitute RUL wired computer workstations, or the range(s) of IP addresses that constitute Rutgers University computers. **We agreed to provide the capability for this restriction option**.

   Access to Rutgers Wireless requires a login using Rutgers netID, and we concluded that we would not include Rutgers Wireless IP addresses in the "Restrict to Rutgers IP" option. In addition, some of the public library computer workstations also require a Rutgers netID to login. Regardless, the "restrict to IP" option should not concern itself with any further authentication requirements in specific settings.
   a. Scenario: Because of copyright considerations, the digital sound recordings of performances or Robert Moevs's compositions can only be provided as "stream-only audio versions on the premises of the Douglass Library, but with no further access." A similar condition has been imposed for the Jazz Oral History sound recordings, for the Institute for Jazz Studies in Dana Library.
   b. Scenario: Email correspondence from dissertation authors, asking to restrict access to a dissertation, sometimes includes a request to "restrict to campus/library", or make it accessible "for educational purposes within the Rutgers library", or provide "limited campus access only" to the dissertation. These users do not intend to be technically precise in their requests – nor should we expect it from them. However, their intention and expectation is clear from the request. These requests can be accommodated by restricting access to a defined IP address range.

2. **Restrict to Rutgers netID users**. A collection or single resource can only be accessed by Rutgers users who are registered with a netID. This access restriction would be used to limit use by a defined "Rutgers user". **We agreed to provide the capability for this restriction option**.

   In our discussions, we acknowledged that certain users who have a Rutgers guest netID (e.g., visiting scholars, adjunct professors, alumni) might not have sufficient credentials to access restricted RUcore resources. We decided to follow the standard Rutgers netID authentication mechanism for this restriction option.
   a. Scenario: Email correspondence from dissertation authors, asking to restrict access to a dissertation, sometimes includes a request to make it "available only to those with a Rutgers netID"; "open to those in the Rutgers' [sic] community"; or require "some qualification (such as RU Id)". The desire is to make it available to students and researchers at Rutgers, with whom the authors have some affinity.

3. **Restrict to defined groups or individuals**. A collection or single resource can only be accessed by a user with a username and password. This will require that we maintain a database of users linked to the collections or resources to which each one has special access. The group can be as small as a group of one, such as the depositor. It could include both Rutgers and non-Rutgers users. When we develop this restriction option, we will have to include a means of credentialing approved users. In most cases, that will be managed by someone designated in the group. Future implementation may consider the use of Shibboleth and becoming a member of InCommon. **We already have begun to provide this restriction option, and will make it available for additional communities**.
   a. Scenario: A research at Rutgers wishes to deposit a research data file. She is concerned about the non-RUcore platform where the data is held, because it is either not sufficiently secure or properly preserved. While she has every expectation that there are no rights or privacy issues that will prevent its distribution through RUresearch, she agrees that, until those issues are resolved, only she will have access to the data. The researcher is a community comprised of one individual. Over time, this community might expand to include research assistants, graduate assistants, or other research collaborators. The lead researcher will approve additions to her community.
   b. Scenario: A group of four co-PIs (two of whom are Rutgers research faculty, and two of whom are affiliated with other universities) need to have reference access to a dataset that was created as a result of grant-funded research among the three institutions. They need to restrict access to the dataset until they have published several articles that are based on the data. The group will comprise a community that has sole access to the dataset. One or more of the group will serve as the community administrator who approves additions to the community.

4. **Restrict from discovery by search engines**. A collection or single resource will be shielded from discovery by Internet search engines. Currently, we have the capability of excluding a collection from the Google site map (for either regular Google or Google Scholar). We considered whether to extend that capability by (1) excluding a single resource, or (2) exclude a collection or a single resource from all Internet search engines.
   a. Scenario: Some ETD authors are disconcerted to find that their work is discoverable by a simple search on their name or key words found in the title. Their expectation is that the work cannot be discovered by Internet search engines. The user requirement is typified

by a graduate student's email: "I found my work online simply by Googling it. I selected that I wanted copyright protection of my work. My point is that I do not even need to be in the Rutgers University system to view my work. I simply Googled the title of my publication and it appears on the RU database."

b. Scenario: This restriction option resembles traditional library catalogs, in which records for library materials can only be found by explicitly conducting a search through the library catalog search interface. If the public user goes to the RUcore search interface itself (or the RUetd search portal), the resource and its description are available to view and download. Similarly, if the public user finds a citation and link (e.g., persistent URL), the resource and its description would be available.

Developing the option to the extent described here would entail considerable effort. **We decided not to pursue this option after taking into consideration the following factors.**

(1) Our metadata – including the persistent URL that links to our RUcore object – can be harvested using OAI-PMH. We also anticipate that we will share our metadata with the Digital Public Library of American (DPLA). Any metadata that is shared in this way can be picked up by Internet search engines. We would have to exclude such resources and collection from being shared with other repositories and databases.

(2) Some will view this restriction option as a step "backwards" from the goal to make more information discoverable on the open Internet.

(3) Our other restriction options are sufficient to protect the intellectual property of authors.

(3) Only a few (two? three?) copyright holders have made such a request. This represents a tiny fraction of contributors.

(4) Such a request could only have been made by someone who lacks a fundamental understanding of Internet behavior. It is not possible to erase the fact of the resource's existence.

5. **Restrict access based on platform type**. This restriction option would allow access only on certain platforms (e.g., mobile devices), or restrict access on those platforms. **We decided not to implement any such restrictions at this time.**

2015-01-09/Ron Jantz & Rhonda Marker