

## Digital Repository User Stories

In the fall of 2014, the CUL Repository Executive Group (RepoExec) created a subgroup to perform a needs assessment for a certain subset of repositories at Cornell: specifically, the “institutional repository”-style systems such as eCommons, DigitalCommons@ILR, Scholarship@CornellLaw, and the pilot currently being implemented for CHE-CALS by Mann.

The subgroup consisted of Jenn Colt, Jim DelRosso, Wendy Kozlowski, Jeff Piestrak, David Ruddy, Simeon Warner, and Mark Williams. The group determined early on that the best format for collecting user needs was through user stories, using the format,

*“As a [user type], I want to [perform task] in order to [achieve goal].”*

For example:

*“As a faculty member, I want to upload my scholarly work to the repository in order to reach a non-academic audience, expand my influence within my academic community, and use it for teaching.”*

With the format agreed upon, we took the following steps to produce the list of 115 user stories contained in this document:

1. Based on our own knowledge of our communities, and outreach to those communities, we constructed basic lists of user stories.
2. We then reached out to other people and constituencies to expand those lists.
3. We revised the user stories so that they were formatted consistently (see above).
4. We applied tags to each user story, using the following list:
  - a. *access/delivery* – how users get resources, which may include issues such as persistent identification, metadata, permissions, embargos, instruction, research, etc.
  - b. *accessibility* – how users interact with the repository system, which can cover various usability issues
  - c. *assessment* – how we evaluate the use of the repository and the resources there in, which can include use stats, metrics, analytics, etc.
  - d. *branding* – how Cornell and associated entities make their association with the resources known
  - e. *discovery* – how users find resources, which may include issues such as persistent identification, metadata, research, custom metadata, linked data, controlled vocabularies
  - f. *ingest* – how resources get into the repository
  - g. *interoperability* – how well the repository interacts with other systems
  - h. *maintenance* – how we manage material within the repository, which may include issues such as system administration, service provision, creation of collections, creation of custom fields

- i. *preservation* – how we care for and keep the resources, which may include issues such as metadata, long term access, understanding, backups, file migration
  - j. *sustainability* – how we care for and keep the system, which may include issues such as workflow, finance, organizational support, creation of collections, creation of custom field
5. After collecting nearly 250 user stories, we divided them up into shorter lists using the tags above. While some stories had multiple tags, this was accounted for in a later step.
6. Each individual list was de-duped and consolidated, and stories that should be re-categorized or removed were marked as such.
7. The lists were brought together for one final round of de-duping, consolidation, re-categorization, and revision; the resulting list of 115 user stories is below.

The list below is separated by category, though it's important to note that any given story was likely associated with multiple tags, even before the consolidation and editing began. We tried to put each story in the best fit category, but there was some subjectivity involved.

The original list of stories can and will be made available as requested.

We hope that these stories will prove a useful rubric for evaluating the various systems that current provide, or are being considered for, repository service at Cornell University.

## Access and Delivery User Stories

1. As a faculty member, I want to access a collection within the repository in order to obtain research materials.
2. As a faculty member, I want to direct students to an item within the repository in order to facilitate my teaching (i.e. desires ability to link to content).
3. As a Cornell faculty member I want to upload data to the repository and associate it with a publication so that I receive attribution for my data as well as my publications (i.e. desires ability to link to external content).
4. As a patron, I want to know which reports/datasets are coming out when (date and time) in order to access latest report/dataset.
5. As a patron or administrator, I want to select/deselect reports/datasets to be emailed, possibly many at one time, when new versions are available.
6. As a DB/IR administrator, I want to mark report/dataset not to be emailed/viewable due to error being discovered.
7. As a DB/IR administrator, I want to manage report/dataset name changes so patrons are able to find historic reports that have a different name from current report.
8. A Cornell faculty member/researcher wants to see clearly what reuse rights they have to files discovered in the repository, so that they know how they can reuse them.
9. As a library staff member, I want to apply or remove an access embargo to an item, because an item author has asked them to do this.
10. As a researcher, I want to stay up to date on research related to a given topic at Cornell, including those in other departments/schools, so my own research remains relevant and innovative. That might include getting notified by email when new materials related to my interests are deposited in one or more repositories.

## Accessibility User Stories

11. As a faculty member, I want to easily collect my publications -- both those within the repository and without -- into a single online access point, in order to raise the profile of my work.
12. As a cross-departmental project team member, I/we need a better place to collect and share project materials, with a range of organization/tagging and access options, including private (internal group use only), University and public. These materials are widely heterogeneous, ranging from raw data to specialized outputs like maps and other data visualizations.
13. As a Cornell faculty member I want to upload data and provide a "cut and paste" formatted citation for people who use it so that I receive proper attribution for my work.
14. As a patron/administrator, I want to view/upload content in multiple formats (e.g. .pdf, .txt, .xls, .zip), including old formats such as Lotus Works.
15. A Cornell faculty member/researcher wants to search the repository and browse through its contents on their phone, to check things on the go.
16. A Cornell faculty member/researcher wants to login remotely to the repository using Shibboleth, to get remote access to Cornell campus-only content.
17. A Cornell faculty member/researcher wants to have some control over how items within their collection sort in collection contents lists, because they are very particular about how their work is presented.

## Assessment user stories

18. As an administrator, I want to be able to use systems such as Symplectic Elements and Activity Insight to properly track faculty scholarly output in the repositories, in order to facilitate personnel decisions.
19. As a library staff member, I want to access useful and comprehensive usage statistics and analytics, including altmetrics, at the repository, collection, author, and item level, in order to assess the results of this service.

20. As an administrator I want comprehensive tracking statistics that can be presented to faculty in a digestible way to increase buy-in for contributing their work.
21. As a Cornell faculty member I would like to be notified if someone downloads any of my items or files so I can follow usage and see what others are most interested in reading.
22. As a DB/IR administrator, I want to see list of subscribers for specified report in order research missing report complaints.

### **Branding user stories**

23. As an administrator, I want to disseminate the aforementioned works in a digital space that is strongly branded, in order to make sure my school is credited for the work.
24. As a faculty member, I want to upload my scholarly work to the repository in order to have a persistent URL that can be used for my personal web page and CV.
25. As a library staff member, I want to create collections which highlight rare, archival, or endangered material, in order to preserve and provide access to this material.
26. As an Extension Administrator or Staff Content Manager, I want easy, well documented and supported workflows that can be easily passed on to others when she/others retire, to maintain continuity.
27. As a faculty member/Extension Educator/Extension Associate, I want to maintain versioning of Extension brochures, FAQs, etc., so only the most current are used/adapted by other CCE Educators in their work.
28. As a Cornell faculty member I want to upload a data file and put a CC-By (or CC0) license on it so that my licensing desires are clearly stated.
29. As a librarian / member of the library staff I want to be able to tell the researchers I'm helping with DMPs that Cornell has a repository system they can use for archiving and sharing results of their research so they can easily meet the needs of funder DMP requirements.

30. A Cornell faculty member/researcher wants to create a collection for their documents, or their organization's documents, and apply some distinctive branding to that collection, to promote their work, or that of their organization.
31. A Cornell faculty member/researcher wants to easily find all the items in the repository that they have authored, by searching on their name (author name authority and/or identification issue), because they keeps their publications in the repository up-to-date and points people tithe repository.
32. A Cornell faculty member/researcher wants to see items they have deposited in the repository appear high up in Google search results, because that's their ticket to stardom.
33. A repository support team member wants to alias (or map) an item from one collection to another (so the item will appear to belong to both collections), because the collection owner has asked them to.

#### **Discovery user stories**

34. As a researcher, I want to discover items in the repository through the library's discovery system and Google, in order to facilitate my easy use of the content.
35. As a researcher, I want to see descriptive metadata for the item whether I come to the item through the repository or directly to the PDF, in order to correctly identify the source of the content.
36. As a researcher, I want to contact the author of a work that interests me, in order to learn more about their research.
37. As a researcher, I want to receive updates when new works are published in a given collection or by a given author, in order to discover more of their work.
38. As a researcher, I want to search within a repository or collection in an effective manner, including searching within subsets of a given collection, in order to facilitate my use of the collection.

39. As an academic researcher, I want to find the works of a particular author in one easily sortable spot so I can see the evolution of their scholarship over time.
40. As an academic researcher, I'm interested in finding the thesis paper of an influential judge or legal thinker to gain insight into how their initial scholarship got started.
41. As a patron, I want to search for report/dataset with keywords or a partial or complete title to locate needed report/dataset.
42. As a patron, I want to browse content by author or subject area to locate needed resources.
43. As a DB/IR administrator, I want to manage subject classifications for ease of patron browsing and searching.
44. As a DB/IR administrator, I want to manage customizable relationship groups in support of patron "see also" option.
45. A Cornell faculty member/researcher wants to search for the dissertations in the repository for which they have been a thesis committee member, to quickly find work by a former student.
46. As a library staff member, I want to make digital items held in trusted external system searchable within our system, in order to provide seamless access to disparate collections.
47. As a researcher/faculty member, I want information that will help me identify other faculty I might partner/collaborate with.
48. As an external user, I want a way to easily find and access research and educational materials from Cornell faculty related to [a particular topic, such as climate change] which I can use to develop stronger research/evidence based programming, grant proposals, community policy, or business practices.
49. As a potential student, I want to learn who at Cornell (on and off campus), has or is engaged in research related to [a particular topic, such as climate change], so I can decide if I would like to pursue studies/research at Cornell.

50. As a potential or current student, I want to learn who at Cornell, has or is engaged in research related to [a particular topic, such as climate change], so I can find an appropriate advisor, choose which courses to enroll in, and identify work/internship opportunities.
51. As a teaching faculty member, I want to identify timely and relevant research suitable for use in my curriculum, and be notified of new research/content.
52. As a patron, I want to view multiple versions of a report/dataset to get both timely and historical information.

### **Ingest user stories**

53. As a faculty member, I want to upload my scholarly work to the repository in order to reach a non-academic audience, expand my influence within my academic community, and use it for teaching.
54. As a library staff member, I want to upload material to the repository on the author's behalf, in order to fulfill a service obligation.
55. As a library staff member, I want to create metadata for items within the repository that includes document title, description, author, document type, status, file locations, subscribable flag, subject area, release frequency, and keywords for searching.
56. As a library staff member, I want to upload large batches of items simultaneously, including descriptive metadata, in order to offer service to authors efficiently.
57. As a library staff member, I want to upload an item to the repository that will only become visible after a specific embargo period has expired, in order to comply with author or publisher requirements.
58. As a library staff member, I want to include supplementary files such as datasets, slideshows, etc. – with DOIs for each, as needed – when uploading a given work, in order to provide a robust image of the work in question.



59. As a library staff member, I want to deliver streaming media from within the repository, in order to deliver such content in the manner expected and ensure continued access to said content.
60. As an administrator I want a platform that can support a variety of media file and text-based formats to showcase and preserve the intellectual output of the school in a more comprehensive way.
61. As a librarian / member of the library staff I want to be able to assign an item to more than one collection.
62. As a librarian / member of the library staff I want to be able to add custom metadata fields for deposits I make on behalf of researchers to allow for better search capabilities.
63. As a DB/IR administrator, I want to upload revised report/dataset due to error in report/document.
64. As a DB/IR administrator, I want to upload new version of report/dataset while retaining previous versions for patron use.
65. As a DB/IR administrator, I want the system to recognize/categorize document type (e.g. historical report still published, expired historical report, single topic dataset, broad topic dataset, report with subtopic, computer file) so that report/dataset is processed and displayed properly.
66. A Cornell faculty member/researcher wants to easily assign a CC license to the files they are depositing, so that users know how the materials can be used.
67. An the repository support team member wants to export a set of records and associated files, manipulate this data set, and reload into the system, because they are aware of a uniform error in a collection.
68. As a Cornell faculty member I would like to associate a researcherID or ORCID with a submission so there is no ambiguity between my work and someone else with the same name.
69. As a student I want to upload to the repository the data that is associated with my embargoed dissertation to meet requirements of a DMP I submitted with my NSF DDIG.

70. As a library staff member I want customizable metadata fields so each project is described in the most accurate way possible.

### **Interoperability user stories**

71. As a faculty member, I want to associate files -- datasets, images, video -- with my scholarly works such that someone accessing the latter also receives access to the former, in order to deliver the fullest picture of my research.

72. As a library staff member, I want to use content and metadata from the repository in other systems and/or web pages easily, in order to deliver such content in the manner best fit to the audience or occasion.

73. As an Organization/Extension Administrator, we need to make resources that were developed for specific projects and geographic areas available to other faculty and educators across NYS.

74. As an Extension Administrator or Staff Content Manager, I want ability to get periodic IR reminders, including when something needs attention (including next ingestion/approval steps), or something that might need to move from a "current" knowledge collection to "Historical".

75. As a researcher, I want to promote/share my work through interactive maps, displaying the range and extent of my work graphically.

76. As an Extension Educator/Associate, I want a place to store a variety of documents, presentations and data sets I've created, that doesn't require an extra step in making available on my own County/Region/Project site (i.e. a list or directory of materials there is automatically updated when placed in a central archive/repository).

77. As a patron, I want to know other reports/datasets that are related (e.g. "see also") to report/dataset being viewed for the purpose finding other related information.

78. As a patron, I want to know previous name(s) and publication date range (first version-last version) of reports/datasets in order to know what historical versions are available and how current more recent version is.

79. As a patron, I want to create a personal account to request email notification when new version of report/dataset is available.
80. As a patron, I want to download report/dataset release calendar to load into own calendar.
81. The operator of library discovery and access services at University YYY wants to harvest useful metadata from the repository via OAI, and continue to do so incrementally, because they have been asked to make Cornell publications discoverable in their repository.

### **Maintenance user stories**

82. As a library staff member, I want to quickly correct errors in uploaded metadata, and even uploaded documents, while leaving a record of my revisions (and possibly the reasons behind them), in order to present correct versions of content and to locate points of error.
83. As a library staff member, I want to receive a quick response and a reasonable resolution to tech support issues, so that service can proceed with minimal interruption.
84. As a Cornell faculty member I want to share on the repository files that are larger than 1GB in a way that still allows users to download them if they want to use them because that's what I work in and I need to share to meet funder requirements.
85. As a DB/IR administrator, I want to login to personal account with access to authorized functions.
86. As a DB/IR administrator, I want a personal account with the ability to change passwords to make them more secure or to retrieve forgotten ones.
87. As a DB/IR administrator, I want to set up or delete accounts for other administrators with various levels of permission (e.g. full administrator, editor).
88. As a DB/IR administrator, I want to see a list of all administrators for ease of administrating accounts.

89. As a DB/IR administrator, I want robust searching and tracking features for existing and newly created patron accounts.
90. As a DB/IR administrator, I want to un-publish a report/dataset if requested to do so by USDA agency.
91. As a DB/IR administrator, I want to add future release dates of reports/datasets to calendar for patron FYI.
92. As a faculty member, I want to recommend a digital collection be built in the repository in order to assist in my research.
93. As a DB/IR administrator, I want to search for administrator accounts using full/partial email or name.
94. As a DB/IR administrator, I want to administrate patron accounts for purpose of closing accounts, viewing subscriptions, changing subscriptions, resetting passwords, changing email address, etc.
95. As a library staff member, I want to create custom metadata fields for certain collections, and allow searching of those fields, in order to best reflect and deliver diverse content.
96. A repository support team member wants to move a collection (move a node to a new parent within tree structure), because the collection owner has asked them to.
97. A repository support team member wants to rename a collection (a node in tree structure), because the collection owner has asked them to.

### **Preservation user stories**

98. As an asset manager/archivist of the current day and the future, I would like the confidence that the digital copy I deposit is stored redundantly, so that when one copy suffers destruction or unwanted changes, I still have other good copies to restore from.
99. As an asset manager/archivist of the current day and the future, I would like to be assured that threats to copies in one repository are not automatically incurred by the

redundant copies (usually satisfied by geographical separation of redundant repository storage).

100. As an asset manager/archivist of the current day and the future, I would like to be assured that the redundant copies are continuously monitored for unwanted changes, and that if/when changes do occur, there is a mechanism by which I can determine the original object and restore it to all copies.

101. As an asset manager/archivist of the current day and the future, I would like to be assured that the assets can be taken out of the repository, and that the assets taken out are the same as those that are put in.

102. As an asset manager/archivist of the future, I would like to know what the intellectual value of the asset is to the institution at the time of deposit, so as to obtain guidance in evaluation of it for current value for the institution.

103. As an asset manager/archivist of the current day and the future, I want a mechanism to obtain a listing or understanding of all the assets that have been deposited, and their current status.

104. As an asset manager/archivist of the current day and the future, I would like the assurance that funding and staffing of the IR have a commitment to longevity of the repository through published plans and/or explicit contract clauses.

105. As a stakeholder in the future, I would like to have files adequately described so I can successfully render them in my contemporary operating environment.

106. As a library staff member, I want to preserve both the content and the metadata within the repository in a robust manner, in order to provide stable, long-term access to this material.

107. As a data curator I would like the ability to upload a new version of an item that has changed, while still allowing old versions to be accessed.

108. As a library staff member I want to preserve the scholarship of the law school by using sustainable links that will not succumb to link rot so I can create a permanent digital archive for the scholarship of my institution.

## **Sustainability User Stories**

109. As a library staff member, I want to create a collection, sometimes with custom metadata fields, in order to house digital material for wider distribution.
110. As a library staff member, I want to trace the list of revisors to determine who uploaded an item to the repository, in order to confirm that proper procedures were followed.
111. As an asset manager/archivist of the current day, I would like the assurance that there is a formal plan for the disposition of assets if the IR should cease.
112. As an asset manager/archivist of the current day and the future, I would like a mechanism to pass ownership/management of the assets to a successor.
113. As a patron, I want to email system provider to report issues and ask questions.
114. As a data curator I would like a mechanism to allow for records scheduling for items that do not need to be stored indefinitely.
115. As a data curator I would like the ability to upload a new version of an item that has changed, while still allowing old versions to be accessed.