

**Columbia University Libraries/Information Services
Strategies for Expanding e-Journal Preservation
Final Report
to the Andrew W. Mellon Foundation**

Principal Investigator: Robert Wolven
Associate University Librarian for
Bibliographic Services and Collection Development
Columbia University

Grant Reference Number: 31300635

Grant Period: November 1, 2013 – July 31, 2015

Reporting Period: November 1, 2013 – July 31, 2015

Summary

Columbia University Libraries/Information Services, in collaboration with the Cornell University Library (jointly referred to as 2CUL) has completed an 18 month project intended to expand significantly the preservation coverage of e-journals and to implement strategies that will sustain the initiative beyond the Project. Some expectations and outcomes have been modified during the course of the project in order to meet the underlying objectives of the proposal. During the course of the project, the project team addressed a number of challenges, and identified areas for future development.

Objectives: Description of the Project and Purpose of the Grant

In October 2013, the Andrew W. Mellon foundation awarded a grant of \$150,000 to 2CUL for use over 18 months (later extended to 21 months) to support the expansion of preservation coverage of e-journals and the identification of appropriate paths to enable libraries to secure the preservation of additional material considered priority content. Methods were developed and tested to facilitate the continued expansion of e-journal preservation through ongoing assessment of priorities and documented practices to encourage and expand community engagement. Preservation strategies best suited to diverse types of e-journal publication were identified, pursued, and promoted. Quantitative analyses were undertaken at the beginning and end of the project period to assess the impact of the project and changes in the state of preservation for a defined set of e-journals.

Deliverables: Expected Outcomes and Benefits

The intended outcomes of the project were three-fold: first, to increase the number of preserved e-journals by 1,500-2,000 through direct action and initiate methods and procedures that would preserve a total of 5,000-7,000 titles within two years; second, to apply qualitative and quantitative analyses to characterize the range of non-preserved titles and, in consultation with other libraries and preservation agencies, to identify high-priority content within this group; third, to develop specific, concrete procedures that would serve to sustain and expand the project's accomplishments following its completion.

Accomplishments: Summary of Activities and Progress Achieved

At the outset of the project two groups were established to advise on priorities and methods and to monitor and evaluate progress: a 2CUL group of librarians representing e-resource management, collection development, preservation, and library technology; and the Collection Development Associate University Librarians in the Ivy Plus consortium (a group of twelve libraries including the eight Ivy League institutions plus Duke, Johns Hopkins, MIT, and the University of Chicago). Recognizing that preservation status is not a binary value but covers a range of conditions, we worked with these groups to define a set of preservation statuses: preserved, protected, in process, and not preserved (see Appendix A.) We adapted our methods in working with different types of content, with the objective of achieving the highest status feasible over the course of the project. We also worked with these groups to validate priorities, including the decision to exclude several categories of e-journals that were either considered low priority for preservation, or as offering little chance of success within the project's duration (see Appendix B). Realizing that the corpus of un-preserved e-journals is large, we wanted to characterize what is considered most valuable from a scholarly perspective and thus poses a greater risk of loss.

Throughout the project we also met and consulted with the major parties currently engaged in e-journal preservation, including CLOCKSS, LOCKSS, Portico, the Keepers Registry, the Directory of Open Access Journals (DOAJ), the Public Knowledge Project (PKP), and the Center for Research Libraries. In addition, careful attention was given to outreach efforts to expand the understanding of issues around preservation for e-journals, as a first step towards expanding the number and range of active participants in this work.

The project focused on five main categories in an effort to develop methods for expanding the coverage of e-journal preservation: content produced or made available by major publishers, small publishers, aggregators, universities acting as publishers, and open access journals. Taken together, these categories gave us a target set of roughly 30,000 active e-journal titles in need of preservation.

Major publishers

Title lists from nine major publishers were evaluated for the preservation coverage of their holdings. The Keepers Registry and data provided by Portico were integral in determining the

preservation status of these titles. Shannon Regan, the Project Librarian, contacted seven publishers with at least one non-preserved title to request further information and to encourage steps to complete preservation. Five of the seven provided detailed responses. In general, these publishers deliberately excluded certain categories from third-party preservation initiatives, mainly: indexing and abstracting titles, newsletters, and continuing education e-journals. Two publishers did not respond to repeated inquiries; however, subsequent analysis revealed that virtually all of their titles in question were later committed to preservation through Portico, though we are unable to say whether our inquiries prompted that action.

Through this analysis two additional factors relevant to the preservation of e-journals from major publishers became evident. First, the extent of preservation coverage for major vendor backfiles varies considerably. Second, the transfer of titles from publisher to publisher can have an adverse effect on the preservation of a title, even if both publishers involved participate in a third party preservation initiative.

Small and mid-size Publishers

In consultation with the Ivy Plus advisory group and with Portico staff, we identified 50 small- to mid-size publishers for detailed analysis and action: 30 already participating in Portico, CLOCKSS, or LOCKSS to some extent, and 20 not participating. Ms Regan again analyzed the preservation status of their e-journals and contacted publishers with non-preserved titles. The obstacles to preservation for small and society presses are in stark contrast to those of major vendors. Small and society publishers face significant cost barriers to participation in third party preservation initiatives. Similarly, most do not have the technological expertise or funds to implement new technologies supporting preservation strategies. Most evident in conversation with small and society presses is a clear lack of understanding as to what digital preservation is and how it may be accomplished. Overwhelmingly, the majority of small and society publishers contacted in conjunction with this project did not participate in any preservation initiative simply because they were not aware of the initiative or did not understand how it worked.

Of the 20 non-participating publishers in this group, two have now joined Portico (with 51 titles preserved or queued for preservation) and two others have been purchased by participating major publishers.

Aggregators

During the second quarter of the project we initiated discussions with two major aggregators of e-journals, EBSCO Information Services and ProQuest, regarding their potential role in preservation and the barriers to preserving the e-journal content they provide. These discussions continued throughout the course of the project. With EBSCO we entered into a deeper collaboration to explore potential business models for supporting a preservation service and to develop workflows for obtaining the rights to preserve the content distributed by the aggregator.

Through this collaboration Ms Regan contacted over 350 publishers. Most of these publishers publish one or two titles. Overall, there was a 30% response rate in which over half of the

respondents agreed to participate in the pilot preservation service. One of the key takeaways from this project is that these small publishers are not resistant to preservation but simply do not know about digital preservation initiatives or the expectations of the library community in regard to the archival status of journals subscribed. Outreach and education have the potential to go a long way in securing preservation, but this is very labor-intensive work, requiring repeated conversation and explanation to preserve a relatively small number of titles.

In order to assess interest in an aggregator-provided preservation service, EBSCO and staff from our project held discussions with EBSCO's Academic Library Advisory Group and with the Ivy Plus Collection Development Group. As of this writing EBSCO has chosen not to pursue offering preservation as a service, but has begun to include the right to preserve content in future agreements with publishers, and to pass that content onto a third-party preservation agency under defined conditions. They are continuing to explore ways in which to act on these agreements so that the content which they distribute is preserved.

Our discussion with Proquest resulted with a different outcome. The Proquest staff involved in the conversations with the 2CUL team were intrigued with the concept of "preservation." They decided to develop and administer a survey to assess the opinions and expectations of member libraries and publishers. They expect to complete and administer this survey later this year. Another related discussion was the pilot work CRL and Proquest have been involved in (within the context of the CRL e-newspaper preservation project) to experiment with a light-TRAC process. The purpose of this concept is to assess publishers/aggregators' technical infrastructures, policies, work flows, and business models in order to assess their preparedness and reliability for performing preservation responsibilities.

Open Access

Freely accessible e-journals comprise the largest, most diverse, and in all likelihood most problematic category for preservation in our target set. They vary widely in importance, in content, and in publication methods and source. There is considerable overlap with other categories addressed in the project, and we excluded from this group open access titles provided through major publishers and aggregators. There are also other active preservation efforts focused on specific groups of open access e-journals. LOCKSS has developed a plug-in that works with the latest version of the Open Journal System (OJS) platform and the Public Knowledge Project is in the process of formalizing their private LOCKSS network. In addition, Portico has a helpful export plugin for OJS users to ensure Portico receives the most accurate files. After consultation with the directors of LOCKSS, the DOAJ, and the PKP program, we decided to avoid overlap with their work and focus on small, independently published titles.

After further conversation with the directors of LOCKSS and CLOCKSS, we decided that the most effective method for immediate action would be to capture the e-journal content through web archiving, with the aim of using either a Private LOCKSS Network or a Fedora repository for long-term preservation. Through its pre-existing web archiving program, Columbia has already preserved 77 e-serials in the Human Rights and Avery Library collections. To test this approach further we selected 77 titles that had been identified as important by selectors within

2CUL. Ms Regan contacted the publishers with details regarding the project and proposed the use of web archiving tools for e-journal preservation. The response was swift and overwhelmingly positive, only one publisher asked to be excluded. Columbia's web archiving team then harvested the content and analyzed the results to identify and resolve issues of scope and quality control. (The collection is available at <https://www.archive-it.org/collections/5921>) We believe this method can easily be extended to a large number of freely accessible e-journals. Cornell is in the process of considering a collaborative project with the Committee on Research Materials on Southeast Asia (CORMOSEA) to focus on the future of open access journals from this SEA region. This nascent vision is considered a component of the SEA newspaper project led by Cornell.

University Publications

A relatively small but growing number of e-journals are published by universities, often through units associated with their libraries. Many are published through the OJS platform noted above, though not all are open access. We decided to devote special attention to this category, recognizing that publishers in this group often have active digital preservation programs whose content is not tracked by Keepers Registry.

We met with the director of Columbia's Center for Digital Research and Scholarship (CDRS) to better understand issues affecting preservation of the 20 journals published by this unit (<http://cdrs.columbia.edu/cdrsmain/projects/#journals>). We examined the websites of a number of other university publishers for statements about archiving practice. Ms Regan contacted a few of the major publishers directly, to inquire about their preservation practices and policies. We found that most of these publishers have not made formal, open preservation commitments, but do plan to preserve their e-journal content.

Impact of the Project

One major objective of the project was to secure the preservation of 5,000 e-journal titles that had not been preserved at the project's outset. It is impossible to say with any precision how many e-journals have been preserved as a direct result of this project. As an example of the complexity involved, one major publisher did not respond to our initiative identifying titles among their online publications that were not preserved; shortly after, however, this publisher deposited over 200 of these titles with Portico. Similarly, Portico engaged directly with some smaller and mid-sized publishers based on our identification of high-priority titles. At least one university library decided to assume archiving responsibility for the journals it publishes after discussion with our project staff. Other titles have made progress towards preservation as a result of our work, but are not yet fully preserved. These include titles from publishers who have now granted archiving rights to an aggregator, and titles in Columbia's web archive; we would consider these to be in "protected" status.

It is even more difficult to measure progress toward a goal of "full" preservation. As we have repeatedly noted, the definition of e-journals (not to mention preservation) is imprecise. We selected e-journals in Columbia's and Cornell's catalogs as a corpus to analyze and focus on. In

the course of our work, however, we encountered and preserved titles that had not been cataloged, while the total number of e-journals added to our catalogs over the past two years appears to exceed the number preserved during the course of the project.

While we can't directly link cause and effect, we were able to analyze the net change in preservation status since the project's beginning for 58,000+ e-journals held by Columbia as of July 2013. This data set was selected for analysis because it had been analyzed by Keepers Registry just prior to the start of the project. In August 2013, 44,889 titles out of 58,556 had not been preserved by any agency tracked by Keepers Registry. At the end of June 2015, 4,043 titles in this set had been preserved by at least one agency. Thus, the percentage of preserved titles increased from 23.3% to 30.2%. The situation improves further when certain exclusion categories are removed from the list. For example, if Chinese, Japanese, and Korean titles are excluded, 34% of the remaining titles are preserved. Similar analysis of three sets of data from Cornell showed an increase of at least 4,570 in the number of preserved titles, with 31.8% of the full set of titles currently preserved.

Beyond the information tracked by Keepers Registry, the impact of the project has been smaller in numbers, but perhaps greater in effect. One major aggregator is now acquiring preservation rights in all of its licenses with publishers. Web archiving has been tested and proven feasible as a means of preserving journals that are unlikely candidates for CLOCKSS and Portico. Recognition of the complexity and diversity of the preservation landscape will increase opportunities for action by more parties, resulting in protection of more of the total e-journal corpus.

Challenges

Several challenges were encountered during the course of the project, and though most had been anticipated, they did require frequent adjustments in the detailed work plan. Perhaps the biggest challenge was simply the time required to explain the purpose of the project, including libraries' expectations and needs regarding preservation of e-journals, to many parties with diverse backgrounds and perspectives. Publishers, editors, and aggregators each had different degrees of awareness of issues, but also different understanding of the meaning of terms such as "preservation" and "archiving." All of those working on the project spent considerable time answering questions. This became somewhat easier as we refined scripts for communicating, developed FAQs, and clarified the language in proposed licenses, but the process of negotiating preservation agreements remains labor intensive.

Adding to this challenge was the fact that preservation is not the highest priority for most of the parties we worked with. Staff at Portico and EBSCO were extremely helpful in providing contact information that helped us reach the right people at the publishers involved; even so, capturing and sustaining the attention of those individuals was difficult. The same proved true in attempting to work with staff at the two aggregators; while our interactions were cordial and informative, it was not always easy to get the attention of those needed to resolve issues.

Changes in the e-journal publishing landscape also required adjustments in project plans. Publishers we planned to work with merged with, or were purchased by others. E-journal titles under investigation transferred to new publishers. New preservation initiatives such as the PKP work on OJS titles were begun. DOAJ expanded the journal description data gathered from the OA publishers, now requesting information about the preservation status of journals listed in the directory. In general, these were positive developments for e-journal preservation, but they did cause a few false starts and course corrections in our work.

Perhaps the most surprising and disturbing challenge was the degree of questioning we encountered within the library community itself regarding the importance of taking action to preserve e-journals. This was expressed as a combination of (in our view, misplaced) confidence that publishers and aggregators can be relied on to archive their own content, plus doubts about the technical and economic reliability of existing third-party preservation agencies. While these views were by no means universal, we did find a number of librarians questioning the costs of preservation, when no major losses have occurred to arouse immediate concern. The reluctance from librarians to aggressively pursue e-journal preservation may be influenced by confusion as to where the responsibility for preservation lies; with publishers, third party agencies, or libraries. Actions to further expand e-journal preservation will bring additional costs. Assessment of the economic impact of this work lies outside the scope of the project, but will be an important area for further attention.

At the outset of the project, one of our goals was working with the BorrowDirect Collection Development group to expand the current model license terms to incorporate preservation-related expectations and requirements. It was during this project that with **funding from the Andrew W. Mellon Foundation** the Center for Research Libraries announced the newly revised model license, incorporating the recent best practices. Although preliminary in nature, the new model license incorporates some licensee rights related to third party archiving of licensed journals.

Project Personnel: Significant Board, Management, and Staff Changes

There were no significant changes in the staffing plan during the course of the project. The timeline for the project was extended by three months, due to delay in filling the project librarian position.

The co-Principal Investigators are Bob Wolven and Oya Y. Rieger from Columbia and Cornell universities, respectively. Joyce McDonough is Columbia's Director of Continuing and Electronic Resources Management. The Project Librarian is Shannon Regan.

Recent Publications, News Articles, or Other Related Materials

Details regarding progress on the project were shared via presentations at several library conferences in 2014-2015.

Oya Y. Rieger organized and spoke on a panel at the Fall 2014 Coalition for Networked Information membership meeting regarding the topic, “e-Journal Archiving: Changing Landscape.”

Shannon Regan presented, “Case Study for Expanding e-Journal Preservation: An in Depth Analysis of De Gruyter’s Acquisition of 67 Berkeley Electronic Press Journals,” at the 10th Annual Electronic Resources & Libraries Conference in Austin, TX on February 24, 2015. <http://academiccommons.columbia.edu/catalog/ac%3A187118>

Shannon also presented, “Strategies for Expanding e-Journal Preservation,” at the 30th Annual NASIG Conference in Arlington, VA on May 28, 2015. In addition, Shannon reprised her NASIG presentation for an audience at the National Library of Medicine on June 15, 2015. <http://academiccommons.columbia.edu/catalog/ac%3A187115>

Oya Y. Rieger & Bob Wolven presented at the Fall 2013 Coalition for Networked Information membership meeting regarding the topic, “Strategies for Expanding e-Journal Preservation” on at Washington, DC.

Next Steps and Recommendations for Further Action

The work completed under this project suggests several promising paths for further action to extend the range and number of preserved e-journals, and to further build community consensus around priorities. The libraries of Cornell and Columbia will pursue some of these actions in the coming months, and will continue working with stakeholders such as LOCKSS, Portico, Keepers Registry, CRL, and the major aggregators to build on this work. The nascent Ivy Plus collection program offers a potential incubator for some initiatives such as web archiving of e-journals, and a platform for publicizing our work. Specific recommendations for future action are included in Appendix C.

Financial Narrative

There were no major variances from the project budget. All expenses were for salaries and benefits for Columbia’s Director of Continuing and Electronic Resources Management at 0.25 FTE for a period of 18 months and for an E-Journal Preservation Librarian for a period of 14-months.

The original budget called for \$2,200 for student assistants to assist with data collection and analysis. This work was completed by other staff and the funding was used to cover minor variances in the salaries and benefits of the two project staff.

The total amount of interest earned during this project was \$7.75.

[Handwritten Signature]

Signature of Principal Investigator

10/28/15

Date

Columbia University Libraries/Information Services
Strategies for Expanding e-Journal Preservation
Final Report
to the Andrew W. Mellon Foundation

APPENDICES

Contents:

Appendix A: Preservation Definitions & Categories

Appendix B: E-Journal Preservation: Quantitative Analysis

Appendix C: Recommendations for Further Action

Appendix D: Final Financial Report

Appendix A: Preservation Definitions & Categories

Criteria for a Fully-Preserved e-Journal Title

For an e-Journal to be considered truly preserved, it must meet the following criteria.

1. Reliability of the Repository

The title has been deposited with a trusted digital repository that:

- Manages long-term archiving based on the OAIS model.
- Has passed TRAC or equivalent evaluation.

2. Access to the Preserved Content

Terms have been established and agreed upon as to how, when, and under what circumstances libraries can discover and access the content and make it available to their patrons.

3. Nature of the Preserved Content

The repository has made available a written definition of the information it preserves (e.g. source files from a publisher, presentation files from a website, ancillary files, etc.).

4. Extent of the Preserved Content

There is confirmation that all retrospective content has actually been deposited, and that deposit is continuing on an on-going basis if the title is still being published.

It is also highly desirable for there to be public notification through one or more sources that the journal has been preserved so that libraries can find it (e.g. a note in the OCLC record, listing by Keepers Registry, etc.).

Levels of Meeting the Criteria

This project will use the following terminology in evaluating the preservation status of titles. In the categories below, where there is a letter, (a) is the best case and quality declines as list goes on. Where there are bullets, no quality judgment is implied.

1. Reliability

a. Secured = meets both criteria.

b. Protected = is in a repository that has a sound track record but that has not met the full requirements (yet). More work is needed, but at the very least there are multiple backups of different kinds, reasonable authenticity and integrity measures, and the repository is sufficiently mature that there is not fear that it may disappear for lack of support any time soon.

c. In process = Efforts are being made, but so far do not come close to meeting the criteria.

d. Not preserved = No one is attempting to preserve this title.

Type of agency managing the digital repository

- Preservation service such as Portico, LOCKSS, Hathi, etc.
- National library
- Academic institution
- Publisher
- Aggregator
- Other

2. Access

- a. Available to all based on trigger events or moving walls.
- b. Available to libraries that had/have subscriptions based on trigger events or moving walls.
- c. Available on-site or within a specific country only (e.g., some national libraries).

3. Nature of holdings

- Presentation files from website (retains look and feel).
- Source files from publisher.
- Includes ancillary files.

4. Extent

- a. Complete retrospective holdings and, for live titles, ongoing of new content.
- b. Substantially complete but with gaps.
- c. Only a small portion deposited.
- d. Nothing deposited.

Appendix B: E-Journal Preservation: Quantitative Analysis

The 2CUL e-Journal Preservation project is based on the recognition that only a minority of the e-journal titles in our collections were preserved as of the project's inception. We have cited figures ranging from 12% to 27%, depending on which data set is used, which preservation agencies are used for comparison, and the time of the analysis. These variables show that no single figure can do justice to the complexity of e-journal preservation.

That complexity is compounded by the wide variety of publications that may be characterized as "e-journals." The term is often, if loosely, used to refer to current (or at least recent) scholarly, peer-reviewed journals published online. The 2CUL analysis, however, included all serials in digital form. It thus embraced many types of publications, with different issues, risks, and priorities involved in their preservation.

The current 2CUL project does not aim to address preservation of all e-journals. To have maximum impact, we are focusing attention and action on only a few categories, and on selected publications within those groups. Beyond that, we plan only to characterize the full data set of digital e-serials in more detail, to propose priorities for action, and to suggest preservation strategies that may be appropriate.

We seek input from the library community, and particularly from our partners in BorrowDirect, on our initial categorization of types of e-journal publication and on our identification of priorities for action.

Analyzing the corpus

Our initial 2CUL analysis was based on e-journal (more precisely, e-serial) titles in our catalogs. The data sets from Columbia and Cornell were not identical, but had high overlap and produced similar results. A separate analysis of data from Duke produced results in the same range. None of these data sets includes all existing e-journals, so we should acknowledge that our project is limited to e-journals "collected" (in some sense) by large research libraries.

On the other hand, our catalogs contain many records for e-journal titles that have not been individually selected by our libraries. They also contain records for historical (non-current) journals whose content is preserved in a variety of ways – print, microfilm, and sometimes digital. The importance of digital preservation is not uniform.

Parsing categories

Ideally, we might use the 2013 Keepers Registry analysis of our e-journal holdings as a starting point, since it provides the most complete analysis of preservation status across multiple agencies. This is probably not feasible because the data provides few clues to the nature of each title (only journal title and ISSNs). We would need to match ISSNs against more complete bibliographic and knowledge base information.

The only practical way to sort ca. 200,000 e-journal titles into categories may be using the source information in the SerialsSolutions knowledge base. Those 200,000 titles are grouped in perhaps 1,000 databases. Information on the provider, together with our own knowledge, makes it fairly easy to categorize each database broadly.

In earlier discussion with Portico, we came up with some possible categories for analysis:

- a) Large publishers (defined by number of titles published)
- b) Small publishers
- c) 3rd party publications (e-journal provided by a publisher other than the original publisher)
- d) Aggregators (agencies that provide e-journals from many publishers)
- e) Open access, freely accessible journals

- f) Databases (collections of e-journal content not organized by journal title and issue)
- g) US Government publications
- h) East Asian journals (mostly in large aggregations)
- i) Historical collections (non-current journals, usually digitized from libraries)
- j) Newsletters, trade publications
- k) Book series (monographs published under a series title)

Broadly speaking, we intend to focus efforts on the first five groups (a-e) and exclude the others (f-k). Some of these categories overlap, and large aggregators in particular will include examples of multiple types. We can begin by agreeing on a hierarchy of exclusion: i.e., first exclude from detailed analysis collections that consist largely of a single category deemed of lower priority (or feasibility) for preservation.

Exclusion Categories

Proposal: exclude from detailed analysis (and from action during the project) collections of e-journals of the following types:

Historical collections: (ca. 50,000 titles)

Description: Collections from either commercial publishers or libraries and non-profit organizations of older journals digitized from print.

Examples: HathiTrust; Gallica; 19th Century UK Periodicals Series

Rationale: Many are already preserved (HathiTrust, Gale collections in Portico); the content is often still preserved in print and/or microform.

East Asian collections: (ca. 11,000 titles)

Description: Large collections of journals (either commercial or open-access) published in China, Japan, and Korea.

Examples: DBPIA; Open J-Gate

Rationale: Securing preservation agreements is likely to be difficult due to language barriers, different legal and publishing environments; individual titles in these collections have not been selected or evaluated.

Databases: (ca. 20,000 titles)

Description: Aggregations of e-journal content designed for searching, without provision for browsing and reading individual titles and issues.

Examples: Factiva; Lexis-Nexis

Rationale: The value of these collections lies in the database as a whole; securing preservation agreements is not likely to be a priority for the database provider; individual titles in these collections have not been selected or evaluated.

US Government Publications: (ca. 21,000 titles)

Description: Serial titles published by US government agencies.

Example: MARCIVE record set.

Rationale: Several other efforts are under way to develop programs for preserving digital government publications.

Newsletters, trade publications: (ca. 13,000 titles)

Description: Collections of minor publications intended for a narrow audience, often providing only a summary of current news within an association or industry.

Examples: Business & Company Resource Center

Rationale: As with the “Databases” category, the value lies in the collection as a whole; most libraries would not choose to collect or preserve the individual titles included.

Book series: (ca. 2,000 titles)

Description: Collections of monographs published under a series title.

Examples: Lecture notes in computer science;

Rationale: The current preservation strategy for titles in these series is to treat them as e-books. Many are deposited in Portico in that form.

Appendix C: Recommendations for Further Action

One of the main objectives of the project was to develop methods and procedures that could be effectively employed to further extend the range and number of preserved e-journals after the project's completion. Given the diversity of content and the range of agents involved in publishing, disseminating, collecting, and preserving e-journal content, future actions will need to match this complexity and involve many different parties.

Outreach and Advocacy: In addition to publishing information about the project and its outcomes and recommendations on the 2CUL website, Columbia and Cornell will seek opportunities to present this work at professional meetings, and to promote continued action through library organizations, archival agencies, publishers/societies, and aggregators, with actions tailored to specific groups for maximum effect.

Major Publishers: E-journals provided by major publishers are in general well covered by existing preservation initiatives. However, most publishers have some titles that are deliberately excluded from preservation, because they are not considered important or suitable for current preservation models. As libraries and licensing agencies negotiate new licenses or renew existing licenses, publishers should be asked to specify any licensed content excluded from the license's provisions for archiving. 2CUL will engage CRL to explore how the recently revised model license can be further enhanced by broadening the archival information section.

Ensuring Continuity: The preservation status of e-journal titles may change as titles move from one publisher to another, either because the new publisher does not have provisions for preservation, or simply through failure to initiate procedures for the newly-acquired content (as was demonstrated in this project.) The Enhanced Transfer Alerting Service maintained by the UKSG (<http://etas.jusp.mimas.ac.uk>) provides information that could be effectively used to monitor such changes. Libraries should work with preservation agencies, the UKSG, and the Keepers Registry to explore means of automatic data transfer, confirmation of preservation renewal, and notification of "dropped" titles.

Aggregators: The project has developed scripts and model license language that aggregators may use in working with publishers to acquire rights to preserve content and pass that content to a recognized preservation agency in case of a trigger event. As yet, however, no mechanisms or procedures exist to effect such a transfer. Columbia and Cornell will continue to work with EBSCO and ProQuest to explore options, but will also seek to engage partners such as Ivy Plus, NERL, and CRL to carry on this work.

University/Library Publishers: Several university and library publishers of e-journals already have provisions for archiving their content through a local preservation repository. University libraries engaged in publishing should develop a consistent approach to preservation, including open declaration of their archiving policies and practice. This work should help to inform, and be informed by, CRL's exploration of a "TRAC light" certification.

Freely Accessible E-Journals: Columbia and Cornell will work with members of the Ivy Plus group of libraries to assess the feasibility and cost of implementing a Private LOCKSS Network to preserve the pilot collection developed in Archive-It, and will pursue this model further in other initiatives currently under discussion around collaboration in web archiving.

Technical Development: As digital formats become more complex and new research methods emerge (e.g., text mining), just-in-case dark archiving solutions will be harder to justify from cost-effectiveness and return-on-investment perspectives. It will be beneficial for the stakeholders to reconsider the current assumptions that underlie significant initiatives such as CLOCKSS, LOCKSS and Portico.

Information Exchange: At present, up-to-date information about preservation status is not included in the systems and knowledge-bases libraries use to manage e-journal content. This inhibits libraries' ability to consider preservation as a factor in collection development and collection management. Keepers Registry has significantly enhanced the ability to query the preservation status of individual titles, but libraries should encourage more systematic exchange of preservation information among preservation agencies, subscription agents, and e-resource management systems.

Setting Priorities: One barrier to effective action has been the sheer number of e-journal titles that are not preserved. The analysis performed during the project helps to identify broad categories that can be used to set priorities, but some of these categories are themselves quite large. More discussion among libraries is needed to build consensus around priorities for action on titles provided through aggregators and on freely-accessible e-journals.

Program/Project Financial Report

The Andrew W. Mellon Foundation
 Scholarly Communications Program
 Financial Reporting Template
 CU Project #: PG006208
 Reference #: 31300635

To be completed by the PI and submitted with the narrative report, or used as the official financial report, if completed and signed by a financial officer.

Organization: The Trustees of Columbia University of the City of New York

Award period: 10/01/2013 -- 07/31/2015
Report period: 02/01/2015 -- 07/31/2015

Income

Grant amount: \$150,000.00
 Interest credited to date: \$7.75
Total Income: \$150,007.75

Expenditures


Categories *	Current Period			Total Project		
	Budget	Actual Expenses	Variance (should be explained in the narrative report)	Budget	Cumulative Expenditures	Balance
Personnel	\$112,783.00	\$40,039.80	\$72,743.20	\$112,783.00	\$113,067.52	-\$284.52
Benefits	\$37,217.00	\$12,951.38	\$24,265.62	\$37,217.00	\$36,940.23	\$276.77
Other Expenses	\$7.75	\$0.00	\$7.75	\$7.75	\$0.00	\$7.75
Totals	\$150,007.75	\$52,991.18	\$97,016.57	\$150,007.75	\$150,007.75	\$0.00

* Please customize or add additional categories to match the approved budget lines in the original proposal

Balance

Total Income: \$150,007.75
 Cumulative Expenditures: \$150,007.75
Total Balance: \$0.00

I certify that to the best of my knowledge all expenditures are for appropriate purposes and in accordance with the approved budget.



 Tamara Hamdam, Executive Director, SPF

10/14/15

 Date

th2271@columbia.edu