Towards Interoperable Preservation Repositories (TIPR)
IMLS Interim Report 4
April 1, 2010 – September 30, 2010

1. Progress towards activity targets

As anticipated in the third interim report, testing activities extended beyond the time allotted in the project proposal, and in fact were incomplete at the end of the original grant project period. The University of Florida applied for and received a no-cost extension, which will allow FCLA and NYU staff to continue and complete the testing. Cornell University declined to continue their participation beyond the original project period.

Hold partner meetings: The final in-person meeting was held April 19-20, 2010 at FCLA. Major topics of discussion were technical details related to completion of the transfer tests, and the practical issue of how to get the widest uptake of the TIPR RXP. The group agreed the best approach would be to get TIPR export integrated into commonly used repository systems, such as DSpace, Fedora, and CDL preservation micro-services, but the best means of accomplishing this was unclear. Although funds were allocated in the project budget for one more partner meeting, the group decided to reallocate the travel funds to support attendance at iPRES, a great opportunity for disseminating information about the project.

Coding: NYU and FCLA finished coding required to ingest TIPR RXPs early in this period. Because Cornell abandoned their plans for an aDORe based repository in favor of one based on Fedora, they did significant coding to implement their new repository. This coding was still in process at the end of the original project period. As a result, coding to implement the function of exporting an RXP from, or importing one to, the Fedora-based repository was never done.

Transfer testing and analysis: Two sets of planned tests were incomplete at the end of the original grant period. In so-called "round robin" tests one repository exports an RXP to another for ingest and subsequent export back to the original repository. Validation should confirm that the chain of custody and integrity of each file and the package as a whole is unbroken. In error tests, one repository sends deliberately mal-formed RXP packages to another, without disclosing the nature of the errors. The receiving repository should be able to discover and identify errors and return this information to the sender. FCLA and NYU have been through several iterations of both types of tests and will complete these in the extension period. If time allows it may be possible to reuse an RXP created from an AIP in Cornell's discontinued aDORe repository to simulate data from a third source.
Draft deliverables: The RXP specification was documented in the form of three complementary METS profiles in June and accepted as registered profiles in July 2010. Finalization of the project website was deferred to the beginning of the project extension.

Dissemination: Partners made an effort to disseminate information about the TIPR project and RXP at major preservation venues as opportunities arose.

- William Kehoe presented "Towards Interoperable Preservation Repositories" at the Fifth International Conference on Open Repositories (OR2010) in Madrid, Spain, July 6-9, 2010.
- Joseph Pawletko gave a well-received presentation on TIPR at the ISO/IEC JTC 1 Study Group on Digital Content Management and Protection meeting in Shenzhen, China, in August 2010.
- The paper, "Towards Interoperable Preservation Repositories (TIPR): The Inter-Repository Service Agreement" by Caplan, Kehoe and Pawletko was accepted for the Seventh International Conference on Preservation of Digital Objects (iPRES 2010) in Vienna, Austria September 19-24 and presented in a poster session at the conference by Franco Lazzarino.

2. Project Findings and Accomplishments

The emphasis of work during this period has been on testing exchanges using the RXP, rather than on major design and development activities. This allowed project participants time to consider the role of an inter-repository service agreement in relation to the RXP and act of transfer. It is well known in the standards world that flexibility and interoperability are often inversely proportional, such that the more flexibility a standard allows, the less it facilitates interoperability. Since the goal of the RXP is to allow repositories to interoperate (in the sense of exchanging AIP content with each other), the format is fairly constrained, having clear requirements and allowing few options to be chosen by the sender and receiver. This makes it technically easy for a repository to implement code to export and/or ingest an RXP, but leaves many details of the transfer process undetermined. The model assumed by the TIPR project is that the details would be worked out by the managers of the participating repositories and specified in an inter-repository service agreement. In a way, the
service agreement would provide the flexibility to communicate context-specific requirements so that the RXP does not have to.

There are formal and de facto standards for service agreements between producers of content and preservation repositories. The Producer-Archive Interface Methodology Abstract Standard (PAIMAS) was developed by the same committee that drafted the Reference Model for an Open Archival Information System and is now an ISO standard (ISO 14721:2006). It outlines 82 steps that producers and repository managers should follow when drafting a service agreement. Although the content of the agreement itself is not specified, it can be inferred from the steps taken to produce it. Into the Archive (Wege ins Archiv) was developed by Germany’s nestor working group on long-term preservation standards, in some measure in reaction to PAIMAS, which was felt to be excessively complex and detailed. The nestor guide is shorter, more practical, and more focused on deliverables.

Both of these standards, however, are focused on agreements between the content owner and the repository; they do not address repository-to-repository arrangements. Some of the specifications apply neatly to both contexts, but others do not. The TIPR project spent some time analyzing the applicability of the standards and developing a model inter-repository service level agreement.

3. Extension period activities

![Figure 2: TIPR Project Year 3 Timeline (no-cost extension)](http://www.loc.gov/standards/mets/METSPrimerRevised.pdf)

As shown in figure 2, the dominant set of activities will consist of completing the transfer tests as originally planned. This is shown as taking place over five months on the timeline only because the FCLA priority for calendar 2010 is implementing a new version of the DAITSS repository system (DAITSS 2), and because the winter holidays are disruptive. The actual number of work hours required should be less than 100.

The RXP format will be documented in a manner similar to the METS Primer document available on the Library of Congress’ website (http://www.loc.gov/standards/mets/METSPrimerRevised.pdf). The RXP documentation will be much shorter, however, because the format is far simpler than METS.

The TIPR website will be reorganized and updated with the most recent information. A final update will be made to include the end of project report to IMLS (not shown on the timeline).
An abstract for a paper about the project was submitted to the IS&T Archiving 2011 conference in November and assuming it is accepted, the full paper will be submitted in February and presented in May in Salt Lake City. Travel funds from the grant will be used for the conference. Other travel for the purpose of dissemination and/or consulting with prospective implementers will be scheduled for some time to be determined during the extension year. It is likely that a representative from the project will be invited to the next scheduled meeting of the ISO Study Group on Digital Content Management and Protection.