Inter-Repository Service Agreement
Agreements Between Sending and Receiving Repositories
29 September 2010

Introduction

Towards Interoperable Preservation Repositories (TIPR) is a project of the Florida Center for Library Automation, the Cornell University Library, and the New York University Library with the goal of enabling the transfer of understandable archival information packages (AIPs) from one preservation repository to another.

Use cases for the transfer of AIPs include:

- diversification for heterogeneity (the owners of valuable content want it stored in multiple, heterogeneous repositories),
- diversification for disaster recovery (two repositories have reciprocal storage agreements, so if one is destroyed, its own content can be restored from the other repository),
- succession (the source repository is ceasing operations and transferring its content to one or more other repositories),
- system migration (the repository is replacing its applications software and must migrate its archived content to the new system).

In repository-to-repository transfer, the source repository creates the RXP from a stored AIP and sends it to the target repository; the target repository transforms the RXP into a SIP and ingests it, creating a new AIP. In order for this to work in the real world, repository-to-repository transfer must have two critical components: a) a standard transfer format, and b) a service agreement between the source and target repositories. The TIPR project developed and tested a set of specifications for a standard Repository eXchange Package (RXP). It also drafted and received feedback on this outline of a generic inter-repository service agreement (IRSA).

IRSA Overview

At a minimum, the service agreement must document the following:

1. details of RXP composition;
2. how the RXP will be transferred from the source to the target repository;
3. actions to be performed by the target repository on receipt of the RXP;
4. archiving and preservation treatment of the ingested RXP by the target repository;
5. rights and permissions agreed upon by the source and target repositories;
6. financial arrangements between source and target repositories;
7. legal aspects of the arrangement.
IRSA Details

1. Details of RXP creation

Where the RXP allows local variation or the provision of locally-defined strings, the specific options and/or terms used should be defined in the IRSA.

1.1. Will the sender’s RXPs in this exchange include the optional rxp.xml.sig file?
1.2. Will the sender’s RXPs include rights information in optional rxp-rights.xml files? If so, how will rights information be represented (i.e. how much of the PREMIS Rights description must the receiver understand)?
1.3. Will the sender’s RXPs include descriptive metadata in optional rxp-dmd.xml files? If so, what metadata scheme(s) will be used? Can the sender provide a schema to validate the xml?
1.4. The RXP requires the METS header of rxp.xml to include an agent element with @ROLE="DISSEMINATOR". What string will be used in the subelement name to identify the sender?
1.5. One representation in each RXP must be labeled “ACTIVE” in the structMap. In this exchange, what does it mean to be the active representation?
1.6. If multiple representations might be included in an RXP, the representations must be ordered. In this exchange, what is the meaning of the order designations?
1.7. Some repositories hold content from a single source, so that the original owner of the content is the same for all RXPs. Many repositories hold content from many sources and keep track of the owner(s) – depositor and/or rightsholder – of each package. How will the receiving repository know, on a package-by-package basis (or in some cases, on a file-by-file basis) the original owner of the content? If this is coded data, what is the mapping to names and contact information?

2. Transfer from source to target

Once the RXP is created, it still must be transferred from the sending to the receiving repository. The mechanism for doing so is outside of the scope of the RXP specification.

2.1. What transfer mechanism will be used (e.g. FTP, OAI, delivery on portable external disk)?
2.2. Will additional packaging be used for serializing RXP contents during transfer (e.g. BagIt)?
2.3. What is the magnitude of the transfer (number of packages, size in bytes, duration)? This should be estimated early in the negotiations, even before exact numbers can be given.
2.4. How will the receiving repository know that the transfer is complete and all packages have been received? What, if anything, will be used as a manifest of RXPs?
2.5. When the transfer will take place? Will it be one-time, occasional, or continuous over a period of time?
2.6. How will the sender and receiver communicate about the transmission? Who is the formal contact on both sides?

3. Actions performed by target repository on receipt

3.1. During transfer, will the target repository acknowledge receipt of each individual RXP, or only of the batch of RXPs? How will the acknowledgement be communicated?
3.2. When will received RXPs be scheduled for ingest? Will there be a quarantine period?
3.3. Will the source repository receive notification of the results of the ingest (or attempted ingest) of each RXP, and if so, what are the details of the notification format and mechanism?
3.4. Will the owner of the content in the RXP receive notification from the receiving repository?
3.5. Can the receiving repository reject an RXP for any reason? If so, how does it indicate rejection?

4. Treatment of the ingested RXPs

Depending on the use case scenario, the sending repository, either as proxy for the content owner (depositor or rightsholder) or in its own right, may have a vital interest in how the receiving repository will treat the ingested packages over time. The relationship of the receiving repository to the sending repository and to the content owners (depositors or rightsholders) must be outlined along with what communication mechanisms are to be used among them.

4.1. What metadata from the received RXP will be retained (events, agents, object information, tipr.xml, etc.)?
4.2. How will metadata in the RXP be retained, in the original format of the source repository, or mapped to the schema used by the target repository?
4.3. What preservation strategies will or might be performed on the content when appropriate (normalization, format migration, etc.)?
4.4. If the RXPs can contain multiple manifestations, will any be discarded/ignored by the receiving repository? Which representations will the receiving repository retain?
4.5. Will different packages receive different types of or levels of treatment, and if so what will determine this?
4.6. Will notification of preservation actions be provided when they occur? If so, who will be notified and how?
4.7. How will ongoing reporting on ingested content be handled? What reports are available? Who will be allowed to request/receive reports, and what are the mechanisms for request and receipt?
4.8. How will future dissemination of ingested content be handled? What DIP format(s) are produced by the target repository? Who will be allowed to request/receive disseminations, when, and with what frequency, and what are the mechanisms for request and receipt?
4.9. Will packages be deleted (withdrawn) from the receiving repository at any time? (For example, if the initial transfer is for the purpose of specialized processing, will the package be deleted from the receiving repository when the processing is complete and the package transferred back to the originator?)

5. Rights and permissions

RXPs may optionally contain an rxp-rights.xml file which is a PREMIS documents describing package level rights. This may not be necessary if all packages have the same rights information which is included in the IRSA. Conversely, rxp-rights.xml may not be sufficient if different files in the package have different rights information. Also, the PREMIS rights entity is deliberately limited to information which is actionable – it includes what actions can be carried out on files and the rights basis for them, but not other obviously useful information like the copyright.
holder. The partners in an exchange should pay particular attention to documenting how rights information will be communicated. If there are restrictions on use, the extent of the receiving repository’s understanding of and adherence to these restrictions should be documented.

5.1. What is the context of rights related to the transferred packages? Are they all subject to the same conditions, or do conditions vary from package to package or even from file to file?
5.2. What are the global rights (that is, conditions pertaining to all packages), if any?
5.3. Do packages contain package-specific, representation-specific, and/or object-specific rights statements? What are the details of encoding?
5.4. What rights and permissions will be respected by the receiving repository? Are there limitations the repository is unable to implement? If so, how will this be documented?
5.5. Are the rights granted to the receiving repository sufficient to allow it to process the transferred material to the extent required?

6. Financial arrangements

Depending on the use case, financial arrangements may or may not be necessary. Because there can be three parties involved in a transfer (the content owner, the sending repository and the receiving repository) it must be clear which party bears costs.

6.1. Will the receiving repository be compensated for expenses involved in the transfer? If so, who will be responsible for payment: the sending repository or the original owners of the content? List the schedule of charges for the initial transfer and ingest and the terms and conditions of payment.
6.2. Will the sending repository or the original owners of the content be billed by the receiving repository for subsequent retention and treatment of the content? List the schedule of charges and the terms and conditions of payment.

7. Legal arrangements

If the receiving repository is to have custody of the transferred content into the future, the responsibilities, liabilities, and rights of the repository and its customers must be laid out.

7.1. What are the standard terms and conditions in effect between the receiving repository and its current depositors? Will this apply to the transferred content?
7.2. Which party is liable for any copyright infringement committed by the receiving repository in ingesting transferred materials?
7.3. Which party is liable for any damage caused by viruses or flaws in the transferred materials?