

## **Islandora as an access system for iRODS managed information packages**

Kilian Amrhein, Zuse Institute Berlin (ZIB), amrhein@zib.de

Marco Klindt, Zuse Institute Berlin (ZIB), klindt@zib.de

### **Session Type**

- Panel
- Presentation

### **Abstract**

Accessing information packages with Islandora is straight forward, albeit not so much when they reside within a federated data management environment. In our case, dissemination information packages live in the Fedora object store for immediate access. The archival information packages are stored safely in a hierarchical storage infrastructure managed by iRODS and are only accessible for administrative and preservation action purposes. We present a data model that supports both use cases utilizing just a single Islandora instance. To integrate with iRODS, we developed an Islandora module to display and deliver data and metadata from the storage location. This solution also allows us to extend the system with further preservation workflow actions that will be required in the future.

### **Conference Themes**

- Supporting Open Scholarship, Open Science, and Cultural Heritage
- Managing Research (and Open) Data
- Integrating with External Systems
- Re-using Repository Content
- Exploring Metrics and Assessment
- Managing Rights
- Developing and Training Staff
- Building the Perfect Repository

### **Keywords**

Islandora, iRODS, systems integration, data management and access

### **Audience**

Our submission might be relevant for data curators, preservation specialists, repository architects and managers, system administrators, developers and those interested in distributed data storage.

# OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

## Background

The extension for our repository provides access to the preservation system with an attached federated storage layer. A new data model enables us to provide different views on the objects. With these solutions we are moving forward the current use cases of an Islandora repository by extending it with an iRODS connector and an administration option for preservation scenarios. In times, where flexible distributed storage is becoming more popular and the demand for preservation as a service rises, this provides a good foundation for an open repository environment.

## Presentation content

The Zuse Institute Berlin (ZIB) [1] is a research institute for applied mathematics and computer science. ZIB also provides long-term preservation services for both in-house projects and external partners. This incorporates the handling of data from cultural heritage institutions as well as research data from multiple disciplines. The entire system is based on the OAIS reference model [2] and provides all the flexibility and safety that is needed to provide trustworthy preservation of such differently structured data. It uses the open source software

Archivematica [3] and Islandora [4], which again is connected to a Fedora repository [5].

It is also build upon a storage infrastructure that comprises both online and offline storage. To have a better manageability of current and upcoming storage resources, as well as to provide an improved and flexible access to data, we use iRODS [6] as data management layer.

The presentation will not be about the ingest process, which is realized with Archivematica, but it will focus on how Islandora is used to present and give access to data managed by the storage system.

As in OAIS we differentiate between Archival Information Packages (AIP) with master files (i.e. TIFF) and Dissemination Information Packages (DIP) with access copies (i.e. JPEG). DIPs live as Fedora Objects in the online storage, whereas AIPs are stored offline with limited access. Both need to be accessed by different user groups. Prime interest of administrators will be an object search and display of submission metadata such as origin of the packages (producer and submission identifiers), contract identifiers or the date of the last fixity check. An external consumer would more likely be interested in the metadata describing the object itself or a representation of it. Furthermore, the latter would normally have limited access rights to the data. Accessing data objects with different roles is straightforward in Islandora with the use of XACML and Drupal's user management. However, there was no interface to iRODS. Thus, we developed a module, which allows us to retrieve data objects such as AIP's and metadata from the storage layer. The link between a Fedora Object and the corresponding AIP is realized with the introduction of a new RDF-triple in the RELS-EXT datastream of the object containing the ID of the AIP. This ID is also stored in the iRODS metadata of the AIP for fast recall of the data. Administrative metadata is extracted upon storage of the AIP and registered with the data object.

This information can be retrieved and displayed under a new storage tab in the Islandora frontend, provided by the new module. A link can be used to download the AIP itself if the necessary rights are granted. The data flow is illustrated in Figure 1.

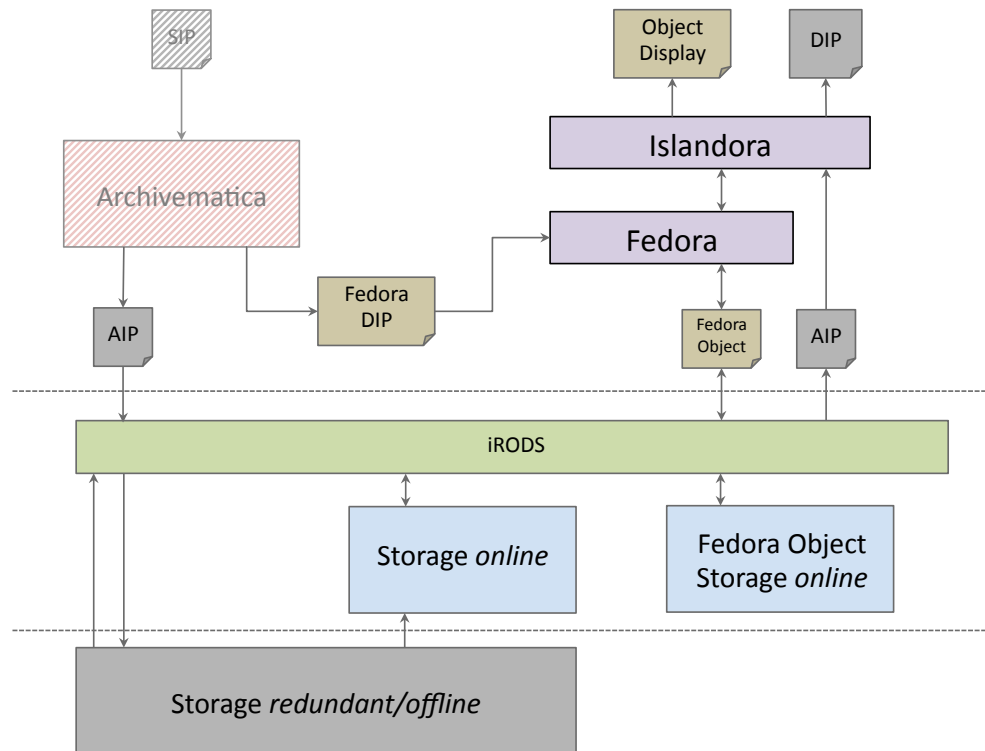


Figure 1: Data management and access structure

To accommodate these two different views of the objects, we also devised a data model making use of the Islandora Compound Object structure, which allows forming groups of intellectual entities by attaching single objects as children to a parent. With this, we now have an administrative access compound object and a content access compound object. From each object within an AIP, i.e. an XML file with descriptive metadata or a TIFF, a single Fedora Object is created containing at least a DC record and PREMIS metadata plus the payload access copies (i.e. XML or JPG) and an optional thumbnail (TN). All these Fedora Objects are attached as children to both compound objects. The contents of the parents' DC datastreams (used for search and display) depend on their types: The administrative object gets its record from the submission information data which is contained in the METS file of the AIP, for the access object the contents of the descriptive XML file (i.e. LIDO) are mapped accordingly. For easy access, the complete METS is also stored within the admin object. The structure is shown in Figure 2.

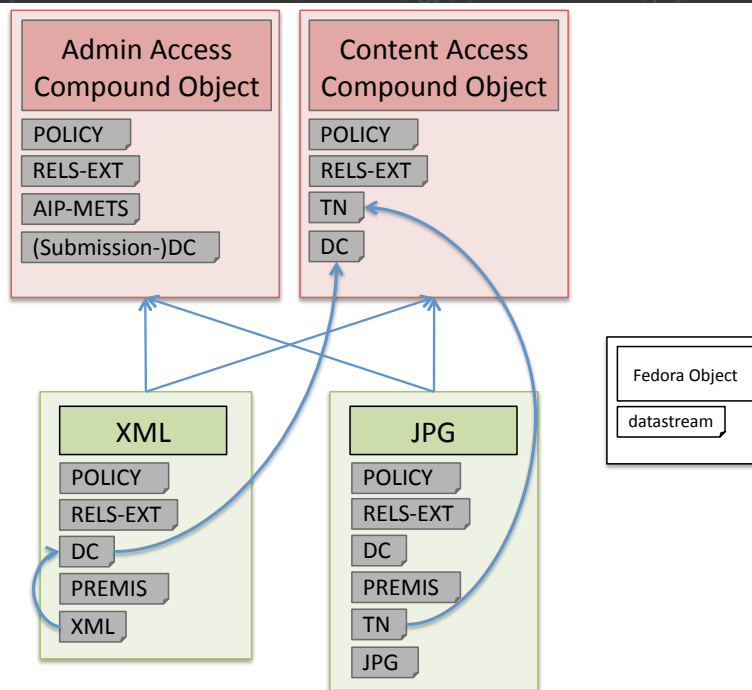


Figure 2: Data model for different preservation system access roles

Depending on the role of the access user (data curator/manager or content user) the objects will be presented differently in Islandora. Specific datastreams such as PREMIS are only accessible for the administrative role. The different compound objects itself are only accessible for the respective role they were created for.

With this setup, it should also be easily possible to implement a mechanism to trigger preservation workflow actions such as migration, fixity checks, or metadata amendment within our curation lifecycle.

## Conclusion

Islandora proved to be flexible enough to serve as an access system for data objects managed by a federated hierarchical storage grid such as iRODS. It also enables us to support administrative and content user views to a preservation item.

## References

- [1] Zuse Institute Berlin (ZIB): <http://www.zib.de/>
- [2] OAIS reference model: <http://public.ccsds.org/publications/archive/650x0m2.pdf>
- [3] Archivematica: <http://www.archivematica.org/>
- [4] Islandora: <http://www.islandora.ca/>
- [5] Fedora: <http://fedora-commons.org/>
- [6] iRODS: <http://irods.org/>