

Digital Preservation with the Islandora Framework at Qatar National Library

Armin Straube
Qatar National Library
PO Box: 5825
Doha – Qatar
Tel. (+974) 4454 1551
astraube@qf.org.qa

Arif Shaon
Qatar National Library
PO Box: 5825
Doha – Qatar
Tel. (+974) 4454 1523
ashaon@qf.org.qa

Mohammed Abo Ouda
Qatar National Library
PO Box: 5825
Doha – Qatar
Tel. (+974) 4454 7190
maboouda@qf.org.qa

ABSTRACT

This poster outlines how Qatar National Library builds a versatile multi-purpose repository that will provide digital preservation solutions to a wide range of national stakeholders and use cases.

Keywords

Digital preservation; Islandora; Institutional repository

1. BACKGROUND

The Qatar National Library (QNL)¹ project was established in November 2012 as a non-profit organization under the umbrella of Qatar Foundation for Education, Science, and Community Development (QF)². The library supports Qatar on its journey from a carbon-based to a knowledge-based economy by providing resources to students, researchers and the community in Qatar. The wider mission of QNL is ‘to spread knowledge, nurture imagination, cultivate creativity, and preserve the nation’s heritage for the future.’

QNL collects and provides access to global knowledge relevant to Qatar and the region. It also collects, preserves and provides access to heritage content and materials. From an operational standpoint QNL has three functions: a national library for Qatar, a university and research library to support education and research at all levels, and a public library to serve the metropolitan area.

All these functions are increasingly fulfilled in a digital way. Addressing issues of digital preservation has therefore become a cornerstone of QNLs operational remit. As a national library, QNL also recognizes an obligation to support other Qatari institutions from the cultural, research and scientific domains as well as other internal and/or external enterprise systems. While many of these institutions host a wealth of digitized and born-digital content, including a variety of research data or output, their preservation over the longer term has so far not been properly addressed. QNL aims to develop digital preservation solutions for both its own needs and for those of partner institutions.

2. QNL DIGITAL PRESERVATION STRATEGY

The QNL digital preservation strategy has been formulated to build a trustworthy digital repository on the basis of established standards in digital preservation and includes the certification of its achievements. The strategy supports a wide range of existing digital collections, including digitised cultural heritage collection³, research data/output generated by Qatari academic/research institutions, and audio/visual materials hosted locally at QNL, as well as various other collections to be established in future.

In general, the QNL preservation strategy is underpinned by a number of guiding principles that serve as benchmarks for the library’s development of its digital preservation efforts and inform its decision making process:

- **Accessibility** - permanent accessibility and usability of all preserved digital content is the principal goal of the QNL digital repository.
- **Integrity** - ensuring the bitstream preservation of archived material is a basic requirement. QNL will take appropriate measures like the regular verification of checksums, multiple storage redundancies and the monitoring and exchanging of storage hardware.
- **Persistent identifiers** - all digital objects will be referenced by a (globally) unique and persistent identifier.
- **Metadata** - QNL will capture technical metadata about all digital objects ingested for preservation and will record information about preservation actions and events using PREMIS.
- **Preservation planning and risk assessment** - all objects ingested in the QNL digital repository will undergo a risk assessment, the result of which will form the basis for decision making on preservation action. The assessment is to be updated and checked regularly to account for technological changes and related economic factors.
- **Standards compliance and trustworthiness** – the QNL digital repository is to be built on the basis of established standards in digital preservation (ISO 14721, ISO 16363, DIN 31644) to ensure longevity and trustworthiness.
- **Development and research via collaboration**- QNL recognises that the complexity and diversity of challenges associated with long term digital preservation is beyond the scope of any single organization. The library will therefore monitor the state of the art in long-term digital preservation and seek to participate in collaboration and research at both national and international levels to facilitate future development of the digital preservation infrastructure as applicable.

Due to the changing nature in the area of digital preservation and the rapid development of services and content acquisition at QNL, the preservation strategy will be reviewed and revised in 2018 at the latest.

¹ <http://www.qnl.qa/>

² <https://www.qf.org.qa/>

³ <http://www.qnl.qa/collections/aihl>

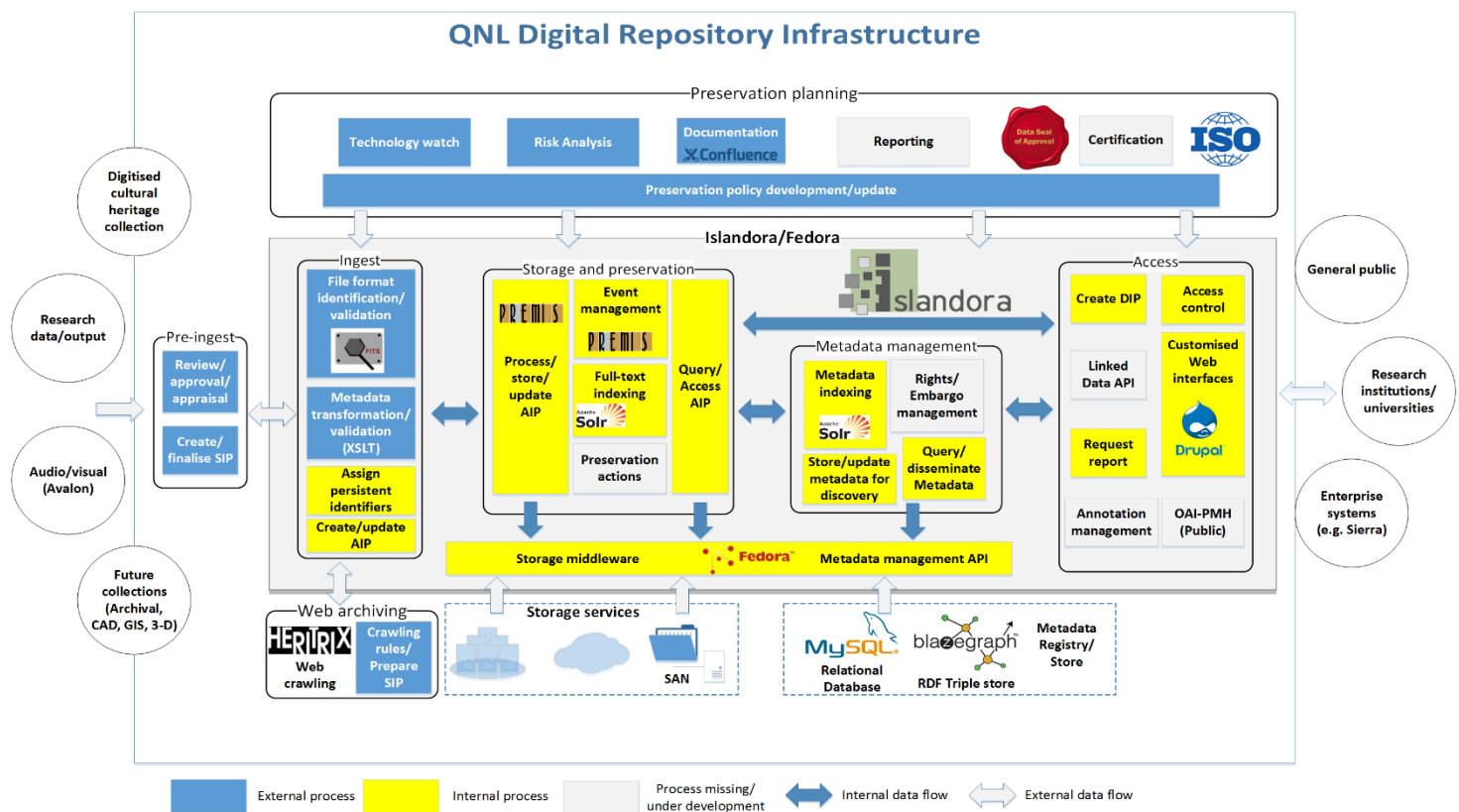


Figure 1: QNL Digital Repository Infrastructure

3. QNL DIGITAL REPOSITORY IMPLEMENTATION

The digital repository of QNL was set up in January 2016 to implement the QNL digital preservation strategy. As illustrated in Figure 1, the overall infrastructure for the QNL digital repository is based on Islandora⁴ (version 7.x-1.6), an open-source digital content management system, integrated with Fedora Commons⁵ (version 3.8.1) as the underlying repository technology. The architecture of Islandora is based on the Drupal⁶ framework that allows different preservation functions of a repository to be developed as Drupal modules, commonly referred to as the Islandora solution packs. The advantages of this modular architecture include customization, further development and integration with third party software.

Islandora provides a range of Drupal modules that support some of the important preservation functions, such as the management of persistent identifiers, the support of PREMIS or the integration of file format identification tools, such as FITS. In addition, the capacity of the Islandora modules can be enhanced by integrating with external preservation solutions – e.g. the Archidora⁷ module that integrates Archivematica⁸ with Islandora. The evaluation/implementation/deployment of these modules is currently under way. In general, Islandora access modules, underpinned by a uniform preservation framework, can be customized to serve a wide range of use cases and be adapted to the need of institutions other than QNL.

While QNL will dedicate its development resources to Islandora, solutions outside this framework can and will also be utilized

where applicable. For example, the audiovisual collection of QNL is made available via the Avalon platform, representing the first deployment of this system in the Middle East.

At present, QNL digital repository mainly stores image based objects (digitized books, maps, photos etc.) in both tiff and jpeg 2000 formats, audio-visual collections in mp4 and wav, and web archives in warc format via integration with Heritrix⁹, an external web crawling tool. Beyond these objects types, the repository is capable of providing bit stream preservation for any digital object and is under development to provide additional support for a wider range of digital objects and different metadata standards.

In addition, descriptive metadata are both stored in the repository and in the library’s Sierra catalogue. Technical systems and workflows are documented in a wiki. QNL has its own dedicated storage infrastructure with tiered storage (hard drives and tape library) and a regular data backup schedule. A policy driven data management is used and multiple redundancies are kept.

4. FUTURE WORK

Future digital collections to be ingested into the QNL digital repository include archival material, GIS and CAD files, 3D scans of museum objects, and databases. The repository will be developed to be scalable to handle increasing volumes of content.

In addition, the library will develop a file format policy, formalizing its current implicit practice, which will enhance the basis of its risk assessment.

QNL aims for certification as a trusted digital repository and will apply for the Data Seal of Approval¹⁰ in 2018 at the latest.

⁴ <http://islandora.ca/>

⁵ <https://wiki.duraspace.org/display/FEDORA38/Fedora+Repository+3.8.1+Release+Notes>

⁶ <https://www.drupal.org/>

⁷ <https://wiki.duraspace.org/display/ISLANDORA715/Archidora>

⁸ <https://www.archivematica.org/en/>

⁹ <https://webarchive.jira.com/wiki/display/Heritrix>

¹⁰ <http://www.datasealofapproval.org/en/>