

## James Hardiman Library

### Digital Scholarship Enablement Strategy

This document outlines the James Hardiman Library's strategy to enable digital scholarship at NUI Galway. The strategy envisages the development of an infrastructure to maximise digital access to, and exploitation of, the Library's archival and special collections, along with the establishment of services and collaborations to support related digital initiatives elsewhere in the University and to participate as a partner in projects locally, nationally and internationally.

At the core of this strategy is a recognition that the Library is only one of many participants in innovative projects focused on collections of digital materials and data on the NUI Galway campus. These include the Moore Institute for the Humanities and Social Studies, Insight @ NUI Galway, the Whitaker Institute for Innovation and Societal Change, Acadamh na hOllscolaíochta Gaeilge and the Huston School of Film and Digital Media. Activities include:

- text and data analysis
- development of supporting analysis tools
- digitisation of archival and other material
- digital humanities research projects
- creation of unique digital collections across a range of media and languages, notably Irish, for use in teaching and research
- deployment of semantic web technologies and linked data to connect with and leverage other collections

Many of these initiatives involve collaboration at and beyond NUI Galway, incorporating a variety of models in terms of personnel, expertise and funding sources. This gives the University an excellent platform for further engagement, potentially across a broader range of activities. Possible deficits locally, however, are that expertise may be unknown and/or unshared, equipment and effort duplicated, long-term sustainability uncertain and opportunities for collaboration missed.

As a central service unit and one of the players in developing digital collections, the James Hardiman Library is developing its own technology and service infrastructure. In doing so, and cognisant of the key role that academic libraries in the United States and internationally have played in enabling digital scholarship, we have consulted widely regarding our proposals. We see great opportunities for the Library and the University through the collocation of archives and research institutes in the Hardiman Research Building, the uniqueness of our archival holdings, the quality and diversity of other players on campus

and the potential for partnerships. With collaboration and ambition the University can excel in digital scholarship as it develops a distinctive strategic plan to 2020.

As regards proposed Library **infrastructure**, the keywords throughout this document include sustainability, scalability, standardisation, compatibility (with the University's IT infrastructure), agility and sharing. We aim to achieve this through a robust IT environment based on standards but with flexible use of evolving, often open-source, technologies and a trusted digital repository at its centre. Other vital elements are systematic application of metadata, a long-term approach to digital preservation, open access where possible, and active collaboration. Opportunities for cost sharing, income generation and commercial use will need to be considered to enable sustainability.

Possible **services and collaborations** may include:

- hosting of digital collections
- publication and discovery
- management of research data
- subscription to identifier services, e.g. Digital Object Identifiers (DOIs), ORCID
- digitisation
- metadata creation and enhancement for linked data use
- long-term preservation
- text and data mining of archival and other collections
- active participation in research projects
- engagement with national and international projects
- joint bids for funding and other awards
- establishment of a suitably equipped digital "maker space" in the Library
- a programme of workshops and training events to develop and share skills

The remainder of this document sets out the proposed building blocks for the Digital Library in terms of technology, metadata, standards, discovery, preservation, rights management, funding, governance and delivery structures. Your input, perspective and advice will continue to be much valued beyond the initial consultation period which concluded in October 2014. The members of the Digital Library Strategy Group can be contacted as below:

Name	Phone	Email
<b>John Cox (chair)</b>	3712	<a href="mailto:john.cox@nuigalway.ie">john.cox@nuigalway.ie</a>
<b>Peter Corrigan</b>	2497	<a href="mailto:peter.corrigan@nuigalway.ie">peter.corrigan@nuigalway.ie</a>
<b>Cillian Joy</b>	5463	<a href="mailto:cillian.joy@nuigalway.ie">cillian.joy@nuigalway.ie</a>
<b>Aisling Keane</b>	4486	<a href="mailto:aisling.keane@nuigalway.ie">aisling.keane@nuigalway.ie</a>
<b>Niall McSweeney</b>	3915	<a href="mailto:niall.mcsweeney@nuigalway.ie">niall.mcsweeney@nuigalway.ie</a>

John Cox, University Librarian, 12 November, 2014

# James Hardiman Library Digital Scholarship Enablement Strategy

## Goals

- Improve and secure the future of NUI Galway’s digital archival and other collections
- Offer services and supporting technologies that are sustainable, scalable, compatible with the University's technology infrastructure, and interoperable with national and international digital library initiatives
- Establish a digital environment that is agile with the ability to respond to change in wider strategies and technology trends.
- Standardise and future proof technology, standards, and related processes.
- Promote and support the scholarly use of digital collections.
- Engage, outreach, innovate, and communicate locally and externally to enable collaboration, partnerships, new services, and content areas.
- Establish a national and international reputation for developing an innovative Digital Library.

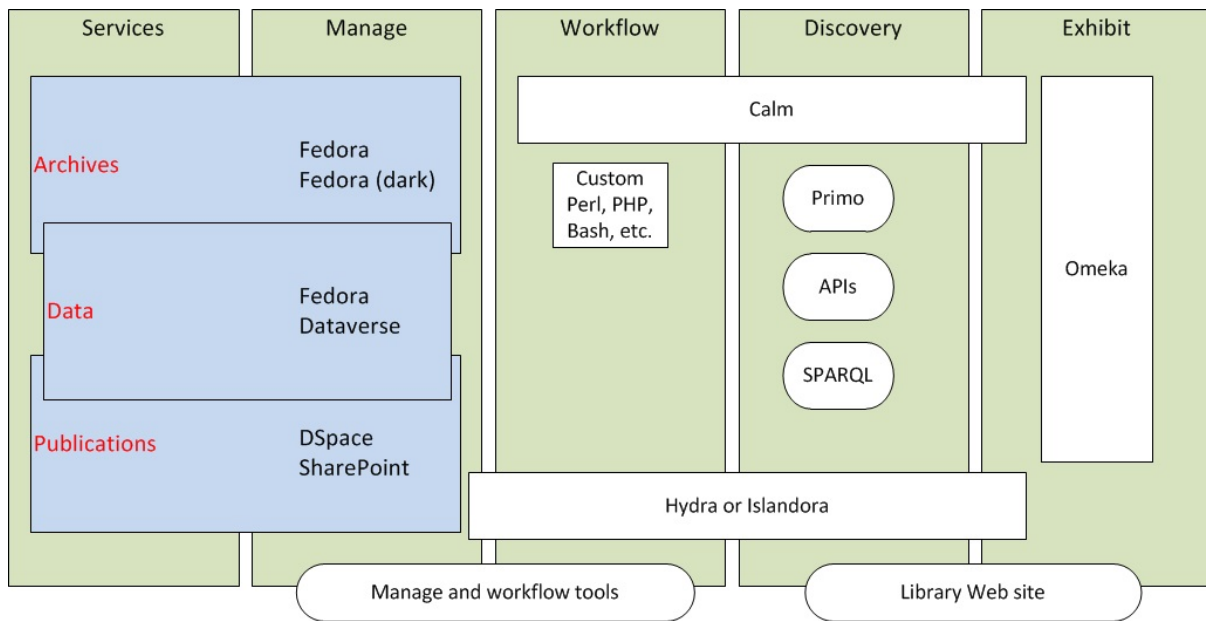
## Technology

The Digital Library technology strategy employs a portfolio of open source and proprietary technologies to leverage unique content to assist research and teaching, while preserving it for long-term access.

### Overview of technology design

The diagram following provides a schematic overview of the proposed design of the Digital Library from a technology perspective.

## Proposed Digital Library Technology Design Overview



## Principles

<b>Interoperability</b>	Interoperability is the primary principle of the Digital Library. Interoperability ensures systems provide and accept information according to defined standards and technology. Information can therefore be migrated, exchanged, or shared between many systems.
<b>Digital preservation and long term access</b>	To facilitate long-term preservation of and access to digital objects, a repository based on the Open Archival Information System (OAIS) reference model will be employed.
<b>Focus on technology to support teaching and research</b>	Supporting teaching and research requires a range of solutions which use the Digital Library for storage, management, control, preservation, and access purposes. The Library's preferred approach is to deploy reusable open source technologies applicable to a range of teaching and learning activities.
<b>Data dissemination</b>	Machine-friendly data dissemination is standard, with a focus on third-party consumption and dissemination via Application Programming Interfaces (APIs).
<b>Access</b>	Will be ubiquitous from the variety of devices. Machine-to-machine and mobile access will be given specific focus.

<b>Align with central University IT resources</b>	The Digital Library aims to align to the maximum with existing solutions, strategies and policies at University level.
<b>Cloud-based solutions</b>	In general, the intention is to maximise cloud solutions to improve efficiencies, reduce cost, and stabilise Digital Library infrastructure.
<b>Open source</b>	Open source solutions are acknowledged as leading the digital library field. The Library embraces, accepts and plans for the related support requirements.
<b>New content areas</b>	Along with services for traditional content areas such as archival material and special collections, the Digital Library will work with new types of content such as research data and geospatial data.
<b>Open access</b>	The Digital Library will enable open access to the fullest extent, but recognising legal and other limits.
<b>Digital Rights Management (DRM)</b>	A variety of DRM scenarios and mechanisms will be supported, ranging from open access to on-campus-only access or other requirements.
<b>Usability</b>	Interfaces for the Digital Library will be user-friendly to support the requirements of end-users and administrators.

## Components

<b>Metadata</b>	<p>Metadata are extensively used as the means to describe, share, search, manage and preserve digital resources, and thus ensure maximum potential for use and re-use. Metadata schema selection, management and administration will be tailored to meet the specific needs of the collection in hand and of its users. Interoperability will be achieved through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) or by packaging one or more descriptions within a Metadata Encoding and Transmission Standard (METS) container.</p> <p>All human-created descriptive metadata will be mediated via an appropriate interface as provided by the repository.</p>
-----------------	--

<p><b>Storage</b></p>	<p>Storage will be technology-agnostic and robust. Key requirements are for storage that is elastic, cost-effective, secure, and sustainable in the long term. Storage for the Digital Library may be a mix of cloud and local solutions, all meeting the same requirements.</p> <p>In the shorter term the University is currently developing a storage strategy and in the longer term storage technology will change. The strategy of the Digital Library is to be flexible to these changing inputs.</p>
<p><b>Discovery</b></p>	<p>The discovery strategy covers both user- and machine-oriented discovery of resources via a number of interfaces.</p> <p>Central to the Digital Library repository strategy is the provision of a generic full-text retrieval facility utilising facets and modern search engine retrieval technologies as a component of the repository itself.</p> <p>For bibliographic and collection-level access records will be provided in OAI-PMH for enterprise discovery tools.</p> <p>For appropriate collections OAI-PMH records may also be provided to Google and other search engines.</p> <p>An OpenSearch API will be available.</p> <p>A SPARQL Protocol and RDF Query Language (SPARQL) endpoint will be available for semantic search and support for linked data.</p>
<p><b>Access management</b></p>	<p>Access management provides controlled and secure access to digital objects. Access management solutions for the Digital Library will be linked with central University IT or national level providers. In general, direct access management is provided by the repository platform. For all audiences, the Digital Library will provide both 1) user authentication (valid user) and 2) authorisation (permitted access). Supporting technology for access management is delivered by University or national level providers in the form of federated technologies which provide access to central user directories.</p>
<p><b>Security and backup</b></p>	<p>Security includes the protection of the Digital Library from unauthorised access, use, change, disclosure and destruction. Backup is an integral part of data management. Testing of restore positions is recognised as important to ensure the integrity of backups.</p> <p>Security and backup for the Digital Library will be informed and guided by the emerging NUI Galway IT strategy.</p>
<p><b>Capacity planning and performance management</b></p>	<p>Capacity planning seeks to ensure provision of sufficient data capacity, in a cost-effective manner, to meet current and future service needs. Performance management means monitoring and providing appropriate processing resources to applications in accordance with service objectives.</p> <p>Capacity planning and performance planning for the Digital Library will be informed and guided by the emerging NUI Galway IT strategy.</p>

## **Standards**

Employing standards future proofs digital resources and processes. The appropriate use of standards for digital materials enables compatibility/interoperability, compliance, consistency, efficiency, and quality. The main advantages of using standards for digital materials are to enable innovation, increased reach, lower cost and decreased time to develop. Standards are required to achieve economies of scale and ensure digital resources are preserved in the long term.

### **Primary standards and principles**

- All Digital Library content and associated digitisation must follow a standard.
- The Digital Library complies with the OAIS (Open Archival Information System) reference model standard and other digital preservation standards and practices as they evolve.
- Digital objects are harvestable using OAI-PMH, the Open Archives Initiative Protocol for Metadata Harvesting.
- A flexible approach aims to select standards most suited to the content type.

### **Metadata standards**

#### **Descriptive metadata**

Descriptive metadata identifies a resource and describes its content. To date, the following are used, according to need:

- General International Standard Archival Description (ISAD[G])
- Encoded Archival Description (EAD)
- Dublin Core
- VRA (Visual Resources Association) Core

#### **Structural metadata**

Structural metadata defines the structure of digital objects. It is used for complex objects with multiple relationships between its components. The Digital Library uses:

- Metadata Encoding and Transmission Standard (METS)

#### **Preservation metadata**

Preservation metadata is designed to maintain access to digital resources and preserve them for the long term

#### **Technical metadata**

Technical metadata focuses on how a digital object was created and other technical attributes, such as its format, storage, location, etc. Technical metadata can be included in preservation metadata (OAIS) or structural metadata (METS).

## **Ownership/rights management metadata**

Ownership/rights management metadata defines usage of the digital object in terms of copyright, access, use, licensing, and reproduction. Ownership/right management metadata is often included in other metadata schemas for descriptive and structural metadata.

## **Discovery**

Discovery means enabling easy access to content and ensuring systems work well together. Discovery is built on and enabled by [standards](#) and [metadata](#). Key elements are system interoperability and the ability to assign unique identifiers to resources.

## **Objectives**

The main objectives for the Digital Library in terms of discovery are:

- The ability to find content quickly, precisely, accurately, and easily
- Multiple presences, same content, using a variety of channels and devices
- Open access to scholarly material to encourage sharing and citing
- Support digital scholarship for teaching and learning by facilitating searching and innovative presentation of content
- Maximise social media technology and trends to encourage “connect and share” and to aid discovery (crowd sourcing)
- Collection-specific promotion and interpretation through blogs and social media
- Link people, materials, and data using standards such as metadata and identifiers
- Third-party consumption to widen discovery

## **Human discovery**

The discovery strategy is to provide both human discovery and machine discovery interfaces. At its centre is the provision of a unified and single search facility utilising facets and modern search engine retrieval technologies. This single search interface provides access to all Digital Library content for users from a variety of devices. While this single search interface provides the standard method for human discovery, other more bespoke discovery interfaces can be used for niche purposes and/or audiences.

## **Machine discovery**

In addition to human discovery, machine-friendly interfaces are used to enable interoperability, wider reach, and to support digital scholarship. For machine discovery, the following objectives apply:

- Provide records in OAI-PMH format for enterprise discovery tools
- Facilitate wide discovery by enabling access to digital material using Application Program Interfaces (APIs). This focus on third-party consumption allows others to



discover Digital Library content. Included are APIs for selected collections and content types as well as the provision of an OpenSearch API

- Digital Library records can also be provided to Google and other search engines, thereby increasing reach
- Linked data is supported using identifiers, while a SPARQL endpoint is employed for semantic search and enabling linked data queries.

## Identifiers

Identifiers facilitate increased discovery of digital resources by supporting both human and machine discovery. Digital Object Identifiers (DOIs) provide identification, attribution, and discovery for digital *resources*. DOIs also provide a mechanism to link data. ORCID identifiers provide identification, attribution, and discovery for *researchers*.

The Digital Library intends to support both DOIs and ORCIDs.

## Digital Preservation

Long-term preservation of digital collections is demanding. The British Library's [Digital Preservation Strategy](#) notes that preserving and providing long-term access to digital content "requires action and intervention throughout the lifecycle, far earlier and more frequently than does our physical collection."

The National Digital Stewardship Alliance (NDSA) publication [Levels of Digital Preservation](#) identifies the following functional areas of preservation:

1. Storage & geographic location
2. File fixity & data integrity
3. Information security
4. Metadata
5. File formats

Focusing on these areas will guide the development of the Digital Library whose performance will be monitored in relation to standards and checklists for trusted digital repositories. Staff skills development will also be planned accordingly.

## Digital Rights Management (DRM)

There are two primary facets of DRM: 1) describing the rights and permissions for access to and usage of digital objects; and 2) enforcing these rights and permissions.

## Description

Digital Library staff will determine and implement the rights and permissions that are associated with our existing digitised materials. A combination of features of the Digital Library repository, a rights expression language (e.g., ODRL, XrML, etc.), and other

appropriate tools will be used to adequately describe these rights and permissions on ingest, and to maintain them over time.

For newly-acquired materials issues such as transfer of copyright, ownership of intellectual property, restrictions on use/copying, licensing and charging, attribution, embargo periods and redaction will be discussed and agreed with their providers. Born-digital donations may present particular issues concerning privacy, intellectual property, and ownership.

### **Enforcement**

Restricting access to digital materials will continue to occur using current and best practices, including hardware and software capabilities. There will be clearly stated terms and conditions of use to which end users must sign up; these will state rights granted, and limits to same, while also stating a requirement for attribution wherever material from the repository is used.

Monitoring and controlling the use and re-use of digital materials present complex technical and legal challenges. These include discovery of user violations and taking appropriate follow-up actions; ensuring clarity of permissions; and licensing of software required to manipulate data sets etc. Furthermore, the legal framework may change, requiring re-assessment of practices.

### **Governance and Delivery Structures**

Three groupings will guide, manage and realise the Digital Library: advisory, governance and operational. An advisory group, representing external stakeholders in particular, will guide development. The governance layer will be primarily composed of Library management and operational team representatives and will focus on:

- Strategy
- Budget
- Policy
- Prioritisation
- Risk management
- Monitoring of progress
- Communication
- Advocacy

Its priority will be to ensure that those projects with highest impact in terms of scholarship, learning and profile are selected and supported following rigorous assessment. The operational team, reporting to a designated manager, will develop and deliver the Digital Library and its associated projects, drawing on skills in:

- Repository management
- Workflow design
- Administration of content
- Assignment of metadata
- Interface design
- API development
- Digitisation
- Data manipulation and curation
- Digital preservation
- Rights and access management
- Project management

Acquiring and developing these skills informs staff recruitment, deployment and development. A focused approach to skills development includes systematic identification and resourcing of opportunities, attendance in person or virtually at relevant training events and conferences, and sharing of learning, contacts, readings and experiments.

## **Finance**

Financial support for the Digital Library encompasses many elements, including:

- Staffing
- Hardware
- Software
- Maintenance
- Storage
- Hosting
- Subscriptions (e.g. Identifiers)
- Membership of relevant organisations
- Service procurement
- Staff development
- Marketing

Some funding is assigned within the Library's operations budget but supplementary sources of funding are essential. There is a particular focus on opportunities to bid, often collaboratively, for funding awards. The approach is to seek out funding awards proactively and to partner where possible with other NUI Galway research institutes or external organisations in project bids. Opportunities to raise income from the Digital Library will be explored, recognising that freely accessible content is the norm and that the cost of administering income collection also needs to be factored into any initiatives in this direction. The development of service level agreements is anticipated, including schedules of charges where required, e.g. to cover shared costs.