

From coexistence to complementarity: the Discovery service and the Archives Hub

In September 2017, Jisc and The National Archives signed a [memorandum of understanding](#) outlining a series of shared goals and areas for cooperation. This included a commitment to “supporting and enhancing discoverability” using our respective online services. In this document, The National Archives and Jisc set out their joint vision for their respective digital archival platforms. Our intention is to collaborate in order to provide the best support for the archives sector and make sure as many archives as possible can publish the best possible collections data online for the benefit of researchers in the UK and globally.

Jisc’s Archives Hub

The Archives Hub has been running as a national aggregation service from its inception and official launch in 2001. It has been developed as a collaborative service, working with archives from across the UK. It continues to build relationships with contributors, providing a busy help desk, a cataloguing tool, and regular workshops around archival cataloguing and best practice. Recently, it introduced an administrative interface, to give contributors more control of their data. All descriptions, at all levels, have persistent identifiers, to help with search engine optimisation and citation, and many have links to digital content.

The Archives Hub originally focussed on archives within higher education but from 2008 it expanded this remit to cover any archives held in the UK that are valuable for research, teaching and learning, and in 2015 it started to cover themed online resources.

The Archives Hub typically has over 600,000 page views per month. It covers over 330 repositories, many providing detailed multi-level descriptions. It takes in new and revised descriptions on a daily basis, using the Encoded Archival Description (EAD) format, and runs processing routines, in order to create more interoperable, reusable data. This enhances the quality of the data and allows the Archives Hub to contribute all of the descriptions to the Archives Portal Europe, Archives Wales and other services, through harvesting and the use of the Elasticsearch API, as well as providing enhanced data back to the contributors. The Archives Hub also provides free local interfaces to higher education repositories.

The National Archives’ Discovery service

Discovery was launched in 2012 as a new service for exploring collections and catalogue data from The National Archives. Between 2013 and 2015 the system became a national collections portal through the addition of indexes from the National Register of Archives, Manorial Documents Register and the catalogue data from the Access to Archives project. The system also mints permanent identifiers (ARCHON repository codes) for all UK archive services.

Discovery contains over 34m descriptions, 9m downloadable documents, over a quarter of a million name authorities and receives over six hundred thousand unique visitors a month, (7.3m sessions in 2016). The service has an extensive coverage, which includes records held by The National Archives, other archives within the public domain and extends to collections held privately. The National Archives' survey of accessions to UK repositories augments the data annually with returns from around 280 services. New data for records held by The National Archives is released to Discovery daily. This growth peaked in the last financial year (2016-2017) with over 1.3 million new or enhanced descriptions made available online.

The newest development to the service is *Manage Your Collections*, a tool launched in 2017 allowing archives to publish ISAD(G) compliant catalogue data directly into Discovery and edit existing entries. In developing tools that enable the ingest of sector data, we will realise the vision of [Archives Unlocked](#) to deliver digital capacity building within the archives sector, increasing impact and in turn access to collections.

Our complementary approach

We want to ensure that researchers – from casual browsers to seasoned academics - have the best possible access to comprehensive information about archives held throughout the UK. We believe that a pluralist approach enables us to explore and implement a variety of perspectives and approaches, to meet different needs. Discovery and the Archives Hub provide different experiences for researchers and enable exploration of interface development, data standards and requirements and ways to engage with end users.

The National Archives has a huge reach and authority as a global cultural heritage organisation and is a sector leader in the UK and widely respected globally. Discovery is enhanced and strengthened by the range of activities performed by The National Archives to curate its own collections and data and also in its role as lead for the archives sector: from cataloguing and digitisation, accreditation and grants to digital preservation and an emerging research agenda. Jisc, in contrast, has vast experience of online services, the use of cutting-edge technology, data processing and manipulation, and working across a variety of domains, so the Archives Hub benefits from Jisc's technical expertise and from operating alongside a whole range of research and learning services.

We can take full advantage of the strengths of both organisations by co-operating and acknowledging our respective skills and areas of expertise.

Interfaces and user experience

Offering different interfaces appeals to different audiences, and it means that The National Archives and Jisc can work on developing our interfaces in different ways. No system works equally well for everyone. Offering different experiences means we can satisfy different end users more effectively. The Archives Hub has an emphasis on data quality and on advanced searching and filtering, concomitant with the Jisc

mission to support higher and further education research and education, it provides descriptions of themed online resources, particularly useful as learning resources, and it offers an interactive map for planning research trips. Discovery brings together a complex range of metadata types (descriptive, structured, unstructured, entities, taxonomies and tagging) and an increasing range of born digital objects which will necessitate new forms of search and presentation. In future, the system will support closer integration of existing research guidance for users alongside search and richer presentation and handling of authority files.

Both The National Archives and Jisc share a belief in evidence based interface development focusing on user needs, testing and audience feedback. Innovative interface design in one service can inform development in the other.

Data

This pluralist approach also works for cataloguing and ingest. Archives vary widely in their ability to supply data and the format in which it can be provided. Multiple systems can more easily meet this range of needs. The Archives Hub has a well-established sophisticated data creation tool, which creates EAD descriptions that can be used in different services. Ingesting content directly from partner archives is a new approach for the Discovery platform and much can be learned from the Archives Hub's huge experience in this area. The Archives Hub can take in EAD exports, so that users of archival management systems can contribute directly through exporting descriptions from their system. There is a particular emphasis on dealing with revised descriptions and ensuring version control. Through the Manage Your Collections tool, Discovery will shortly support highly flexible upload, allowing contributors to dynamically match their data schemas to the full range of fields present in Discovery.

The Archives Hub emphasises data standards and interoperability; Discovery is looking to ensure that repositories with minimal data descriptions are not excluded from being part of a national aggregation. The Hub has developed expertise in the use and benefits of EAD; Manage Your Collections is developing ingest through Excel in the first instance. The Hub requires data to be submitted to a central store; Discovery is exploring options for indexing catalogues 'in situ' on host sites. The Hub data is automatically provided to the Archives Portal Europe; Discovery provides the benefits of cross-searching alongside The National Archives own massive collections.

This variety allows repositories to assess a range of options and consider which approach best suits their needs. Repositories can choose to contribute to Discovery or the Archives Hub or to both platforms.

Whilst there may be some argument for bringing these different facets together into one aggregation service, the reality is that with two expert and experienced players in the space, we can specialise and put effort into different approaches and innovations. At the same time, by working together more closely, we can look at how to align the services, and we can assess the potential for more integration over time.

Outreach and support

Both organisations work hard on raising awareness, outreach and support. We can do more together to help promote archival collections and bring them to the attention of a wider audience. The National Archives has published its position on [digital cataloguing practice](#) and will promote best practice in the sector, whilst also working with The Pilgrim Trust to provide the only funding stream available in the UK dedicated to cataloguing and unlocking archives through [Archives Revealed](#). The Archives Hub has run regular hands-on workshops on cataloguing and best practice in online discovery for many years, and is now developing workshops relating to exporting EAD from archival systems.

The National Archives will work together with Jisc to modernise descriptive practice for the digital age and promote enhanced data quality from archival services and from the system vendors providing the most widely used cataloguing systems.

Sharing

We are examining possible approaches to connectivity between the two services. Such approaches may take a variety of forms. Both services are fully open to search engines and offer citable URIs. Both services offer open licensing and APIs, enabling data sharing in principle. Our challenge is to work together on implementing and developing data standards and practical ways to successfully bring diverse datasets together.

We are committed to a productive, fruitful and active ongoing relationship and we are confident that working to our strengths and working together will bring substantial benefits to UK archives and to our end users.