

Appendix N

Discovery (<http://discovery.ac.uk>)

The JISC -funded Discovery programme was launched in May 2011 to create 'a metadata ecology' to support better access to vital collections data in libraries, archives and museums and facilitate new services for UK education and research.

Our work will continue to the end of 2012 and is focused on advocating open data, reducing technical and licensing barriers, providing information, advice and training, and supporting exemplars.

The Discovery programme takes forward the Vision of the earlier JISC and RLUK Resource Discovery Taskforce (RDTF) which has been working with partners from the libraries, archives and museums since 2010.

Vision

The Resource Discovery Taskforce (RDTF) Vision is about making resources more discoverable both by people and machines. In doing so, we can make our data work harder, integrating and combining it in new ways that can add value for researchers, teachers, and managers of information assets within libraries, archives and museums:

"UK researchers and students will have easy, flexible, and ongoing access to content and services through a collaborative, aggregated and integrated resource discovery and delivery framework which is comprehensive, open, and sustainable."

Working to enable this Vision, JISC and its partners and the Discovery team have now responded by committing to a programme of activity. In this context, 2011 is about building critical mass through opening up data, providing examples and communicating outcomes.

Aims

Discovery will....

1. Clearly position and define the benefits of Discovery to research and education at the local and national level
2. Improve the discoverability of UK library, archives and museum content
3. Drive a shift in ethos to 'open' in institutions, services and funding bodies
4. Improve the quality and sustainability of new and existing resource discovery infrastructure
5. Be understood, endorsed and promoted by key stakeholders within the library, archives, and museums sector and beyond

Targets:

1. Progress embedding of the technical, licensing and metadata principles
2. Drive innovation and sustainable, benefits-led reuse of LAM open metadata
3. Identify and establish core efficiencies in dataflow and aggregation that can be achieved by key shared UK bibliographic data services
4. Establish open licenses for JISC library and archives service metadata and other key UK LAM aggregations
5. Develop demonstration exemplars of what is possible, strengthening the business case for open data and also identifying issues for sustainability
6. Open up and make discoverable important but hidden collections
7. Demonstrate and support approaches to inaccessible metadata and where no metadata exists
8. Persuade funding bodies and vendors to support the key principles

9. Engage with related initiatives to ensure that the approaches recommended in Discovery are compatible with relevant work occurring elsewhere.
10. Work with related JISC initiatives to explore how they can be integrated into the Discovery framework

Open Metadata – Developing the Business Case

The idea

Discovery is advocating the adoption of open metadata for the furtherance of scholarship and innovation. That's an important business case for UK plc – but what could that mean for individual cultural and educational institutions and agencies? Is there a local business case for open data? The Discovery principles, launched by the Resource Discovery Task Force on 26 May 2011, propose that '**Open metadata** creates the opportunity for enhancing impact through the release of descriptive data about library, archival and museum resources. It allows such data to be made freely available and innovatively reused to serve researchers, teachers, students, service providers and the wider community in the UK and internationally.'

There is no doubt that libraries, archives and museums exist in an environment where they, their stakeholders and partners share an interest in exposing and repurposing metadata. The growing momentum around open data indicates that as a community and as a country we are beginning collectively to agree that open data provides a platform on which innovation and value generation can flourish. To achieve the vision we need to publish more metadata openly and unambiguously and so reduce barriers and promote a virtuous flow of information.

Such approaches can be catalysts in realising new benefits by making existing data work harder. Much metadata has been created through public funding, so this work is rooted in increasing the return on public investment by fuelling innovation, enhancing learning and thus creating new knowledge. This ecosystem should encourage more entities and individuals to enter into the supply chain and add value to that content in new and surprising ways. It should also enable curators and consumers break out of the walled gardens (even silos) that have historically dominated resource discovery. Consequently, as Professor David Baker, Chair of the RDTF and JISC Deputy Chair states, 'we should be better equipped to serve UK educators and researchers to excel in their work by increasing visibility of and access to relevant content.'

We must however recognise significant differences across the spectrum of libraries, archives and museums. Whilst business drivers vary and use cases are distinct, our overarching aim across these sectors is the same - we want our collections to be instrumental in teaching, learning and research. Given the paradigms of the web, that aim is most likely to be achieved if those collections are discoverable through popular search engines as well as through specialised services and aggregations, and if they can be exposed through social platforms ranging from scholarly reference management systems to Facebook and Twitter.

Change is already happening

This belief in the potential of 'open' is already motivating UK libraries, archives, and museums to open up their metadata for re-use. The **British Library** has recently released a new sample from the British National Bibliography under a Creative Commons CC0 license. Other adopters include the University of Cambridge, currently working with the Open Knowledge Foundation and with OCLC in projects to make large numbers of bibliographic records freely available. In the museums sector, Culture Grid is opening up UK collections information and sharing data under the Creative Commons CC0 open licence. The Archives Hub, representing over 160 archival repositories across the UK, has recently released a substantial subset of this aggregation for reuse, again under

CC0. Meanwhile, The National Archives has been instrumental in driving the <http://open.data.gov> site, and has recently launched an open programming interface for legislation.gov.uk. These are just a few of the UK institutions that have identified the case for opening up data for reuse, inviting serendipity and actively investigating the possibilities that emerge as a result. These open initiatives compliment a growing number of European releases from cultural and research organisations such as CERN, the Amsterdam Museum and the Swedish National Library.

More than a call to arms

We are mindful that, in addition to advocacy, we need to better articulate the risks and barriers in this area to support decision-making and help data providers manage the transition to open licensing. **Discovery** therefore seeks to help libraries, archives and museums, large and small, not only to identify more fully the possibilities for open data but also to overcome cultural, legal and technical barriers to releasing their data in this way.

Discovery and its partners, such as RLUK and the Collections Trust, are conducting work across the sectors to understand perceptions of open and what this means on the ground and to develop targeted informational and decision support tools. For example the JISC has made available the Open Bibliographic Data Guide website which details around twenty use cases to help organisations deal with issues such as how to license data and the practical implications. Our next task is to produce similar guidelines with our partners that are relevant across the sectors, and over the next year we will be working closely with the sectors to accomplish this.

Moving beyond rhetoric

A leap of faith, believing in 'open' as part of a greater cultural change and as a national differentiator is not enough. We know that 'open' does not mean 'free' – there will be costs associated with opening data, and across the sectors people are naturally asking what the benefits of opening up their data is to them. How does this help achieve institutional objectives? Why is this worth their investment?

In tough economic times, institutions need a clear business case. Recent interviews with senior decision-makers indicated that 'the business case for open data needs to be clearly aligned with institutional mission' and warned that 'managers who can themselves see clear benefits of open data for the quality of service provision have difficulty articulating these benefits in terms of a valid business case to present to university management.'

As we engage the wider community beyond early adopters, interviews with practitioners and managers alike emphasised that we need more resonant examples and illustrations of what is made possible by open data. Only then will it be possible for professionals to appreciate the scope and scale of the opportunities and to convince managers and funders of the value of the approach. In the coming months **Discovery** will therefore be working with existing projects and interested institutions and agencies to extract the key lessons – strengths and opportunities, not forgetting and weaknesses and threats.

Recasting the value chain

The question of the economic value of open data and particularly 'who profits' is a complex and potentially vexed. Relinquishing control over data means that wealth might be generated outside the public bodies generating the data. However current thinking emphasises that to address the challenges of the information age, we need to rethink the perception that we are exclusively gatekeepers to or guardians of this content. To accept this as a possibility, even as a necessary step, represents a fundamental shift in ethos, a disruption that will not sit comfortably with all.

We conclude therefore by highlighting some of the local benefits that might arise from a commitment to the release of open metadata covering library, archival or museum collections. Here is a short list of five considerations out of many more:

- Libraries can enhance support for efficiencies such as collaborative cataloguing and collection management
- Memory institutions can combine information to provide a more complete set of signposts to support a richer range of narratives and user quests
- Any special collection can become more discoverable and therefore more widely used
- Aggregators can be enabled to work more innovatively to promote exposure of contributing collections
- The wider community of developers, of finding aid authors and of narrators can be leveraged as co-creators to benefit access and articulation in both planned and serendipitous ways
- Overall, institutions can focus their efforts on adding service value and providing authentic raw material, rather than on preserving the dikes and halting the waves

This is unlikely to mean freeing up everything – but it should precipitate a step change in core thinking, where the first business case question becomes 'Why not?' – that, in the words of the **Discovery** principles, 'institutions and agencies should proceed on the presumption that their metadata is by default made freely available for use and reuse, unless explicitly precluded by third party rights or licences'.

*The **Discovery** initiative and this movement more broadly is about embracing and facilitating the growth of new business models, not only rethinking our value proposition but also reflecting on our very purpose. We hope you will join us, not in blind pursuit of an ideal but rather by contributing to the community dialogue about rationale and business case and consequently to the shared reservoir of open metadata.*