Against the Grain

Volume 25 | Issue 5 Article 8

2013

Lessons Learned: Jisc's Experience in Acquiring Multimedia Resources on Behalf of UK Academic Libraries

Lorraine Estelle
Jisc Collections, l.estelle@jisc-collections.ac.uk

DOI: http://dx.doi.org/10.7771/2380-176X.6604

Lessons Learned: Jisc's Experience in Acquiring Multimedia Resources on Behalf of UK Academic Libraries

by **Lorraine Estelle** (Chief Executive Officer, Jisc Collections) <l.estelle@jisc-collections.ac.uk>

Historically, **Jisc** stood for "Joint Information Systems Committee," but over the last decade it has evolved into a charity simply known as **Jisc** with the mission to enable people in higher education, further education and skills in the UK to perform at the forefront of international practice by exploiting fully the possibilities of modern digital empowerment, content and connectivity.

Since 2001, **Jisc** has licensed moving images and still images for its community and has also funded the digitisation of many special collections of news film and still images. **Jisc** now provides 3,600 hours of film, 55,000 still images, and 50 hours of classical music, all copyright-cleared to the UK academic community.

In those early days of multimedia acquisition some great collections were acquired, including *Logic Lane*, a series of films tracing the development of philosophy at **Oxford University** from the 1930s to the early 1970s and featuring such eminent figures as **Iris Murdoch** and **Sir Isaiah Berlin**. However, errors were made too, including licensing collections from commercial providers who demanded what turned out to be unsustainable annual licence fees, and licensing collections for which there was little demand.

In 2006 **Jisc** took stock and published *Digital Images in Education: Realising the Vision*, a book that was instrumental in mapping out the future for this area of activity. The vision explained in the book was to:

> …provide the UK education community with long-term access to the digital image resources that it needs, in a variety of convenient, flexible and easy-to-use ways. Ideally, provision should be free at the point of use; comply with common open standards, cover the broadest range of possible subject areas; have copyright clarity, be sustainable; and support maximum usage at all levels of teaching, learning and research. (Williams, 2007)

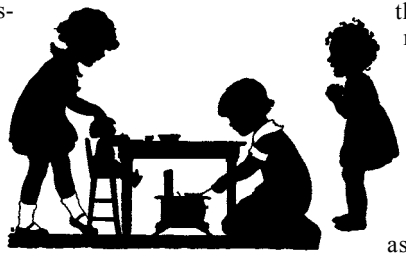
In 2008, **Jisc** was fortunate enough to receive funding that provided a significant building block in realising this vision. Funding to acquire a large collection of images and moving images relevant to many areas of the curriculum in further education and the subjects studied in higher education. These collections would provide pictorial evidence of world events over the past 25 years — a period specifically chosen because licensing costs usually make more recent images unaffordable for most academic institutions.

Based on previous lessons learned, we decided that rather than engaging in individual negotiations, the procurement would be done through a tender process. Not only would such a process provide transparency, it would also ensure that the vendors clearly understood the requirements for compliance at the outset.

The main requirements for compliance were:

- All images must be copyright cleared for educational use. Once an institution has agreed to the terms of a licence, all staff and all students of the institution must be able to use the images (in conjunction with educational activities such as teaching or research) freely and without further authorisation.
- The images acquired must be supported by open metadata, which should include geospatial tagging. The moving images must further be supported by encodings.
- Ideally a perpetual licence, or if this could not be granted, a licence for a minimum of 25 years.

Of these, the requirement to provide supporting metadata was the most contentious and challenging to achieve. At the briefing session to support the tender process the vendors questioned the need for metadata and encodings. Some of the vendors argued that they could provide many more images and moving images, if they did not need to specially create metadata to meet **Jisc's** requirements. The procurement team had to continually iterate that supplying metadata is of primary importance — there is no point in having excellent images and films if users cannot find them. Rich metadata will ensure resource discovery!



As mentioned above, one of the lessons **Jisc** had previously learned was that it was all too easy to be "supply-driven" in this area and licence material that would be little used. An objective identified in *Digital Images in Education: Realising the Vision* was:

> Plans are required that help the community move away from provider-led, controlled management of resources and towards an open, sharing culture wherein development of resources is led by direct user involvement and genuine needs. (Williams, 2007)

Thus, full participation from the academic community was needed to evaluate the wealth of bids received. A call for volunteers was issued and panels of experts from higher and further education assembled. The panel members included librarians, teachers from a wide range of disciplines, learning resource managers, geospatial experts, and metadata experts.

The panel members were rigorous in rejecting collections that failed to comply with the requirements stated in the tender, or that they considered would have small value in education or research.

The result of the tender was awards to the following vendors:

Moving images:

AP Moving Image
Endeavour Getty
ITN Source

Still images:

PYMCA
Endeavour Getty
GoveEd Communications

The tender process was completed but the fun had barely started! The successful procurement meant that some 80,000 still and moving images were contracted to be delivered between April and December 2009. That success brought with it a problem because **Jisc** had to find some way to evaluate those images as they arrived: the tender process simply evaluated a small number of samples. This meant that over an eight-month period, a team of experienced evaluators assessed approximately 6,500 images and their associated metadata each month (200 images per day)!

Jisc also had a logistical problem because the evaluators would need to look at images alongside the metadata and approve or reject them or send them back for correction. Hence a team of postgraduate expert evaluators had to be contracted and a "loading bay" had to be built. The loading bay enabled the successful bidders to directly upload the material in batches as it became available and eliminated the need to transfer large files to the evaluators.

Jisc held a training workshop to explain to the successful bidders what was required of them for the creation of metadata and encodings for the moving images. Nonetheless, obtaining consistent and correct metadata was challenging. The project drew on images and moving images from commercial collections. In a commercial context, images are sold typically for advertising purposes and are tagged as "woman," "night," "smiling," for example; all perfectly helpful for advertising purposes but not for education.

continued on page 18

Lessons Learned
from page 16

The requirement for geospatial tagging was also problematic. For example, the following caption explains that the painting in question was once in Florence, but the aim was to have dates and places held in separate fields so users would be able to search by dates and date ranges rather than through a simple string-based search.

> This is one of two panels that were part of the predella that forms the lower edge of the large altarpiece of **Veneziano's** "St. Lucy Altarpiece" (c.1442-48). Originally in the church of St. Lucia dei Magnoli in Florence, the altarpiece appears to have been dismantled by 1816.

A large number of images and moving images were rejected at the initial evaluation because of spelling errors in the encodings or metadata. This problem particularly applied to the "rushes" (the never-before seen unedited footage from which news broadcasts are selected), which **Jisc** had encouraged the vendors to provide. It should also be remembered that commercial providers usually compile metadata for internal use, rather than for publication, and so most of the encodings and metadata supplied had not been through any form of editorial review.

The logistical and metadata problems overcome, the project produced more than 500 hours of film clips — from **Gorbachev's** accession to power in the Soviet Union in 1985 to the financial crisis of 2009, and including powerful raw footage of the 9/11 attacks as well as coverage of key issues such as deforestation and global warming. All told, a large and diverse collection of over 56,000 photographs to support teaching and lifelong learning was developed in the areas of history, social sciences, science, art and creative industries, and geography.

These collections were and continue to be delivered to the UK academic community through a service called **Jisc** MediaHub, which provides a single point of access and enables users to search and link out to other external media collections such as the Open Video Project, Wellcome Images, ADS, ARKive, and the First World War Poetry Archive.

In summary, although **Jisc** usually negotiates with vendors on behalf of libraries, in the area of media resources we recommend a tender process, not least because this ensures a very clear definition of requirements and evaluation process. Evaluation by educational experts is essential in building collections that will be of value in research and teaching and provide a long-term return on investment. Licenses in perpetuity — or for at least a very long term, are essential, because it is impossible to sustain annual subscription fees in an uncertain economic climate. Finally, metadata is king! However interesting or informative an image, it is useless if it cannot be found.

Reinventing the Methods Journal: Increasing Reproducibility with Video Journals

by **Kira Henderson** (Deputy Director of Journal Development, JoVE)
<kira.henderson@jove.com>

The way science journals present research must be rehabilitated or risk becoming obsolete, causing foreseeable negative consequences to research funding and productivity. Researchers are dealing with ever-increasing complexities, and as techniques and solutions become more involved, so too does the task of describing them. Unfortunately, simply explaining a technique with text does not always paint a clear enough picture.

Scientific publishing has followed essentially the same model since the original scientific journal was published in the mid-seventeenth century. Thanks to advances in technology, we have seen some minor improvements such as the addition of color printing and better dissemination and search functionality through online cataloging. But what has actually changed? In truth, not all that much. Articles are still published as text heavy-tomes with the occasional photograph or chart to demonstrate a point.

Dr. John Ioannidis, the C.F. Rehnborg Chair in Disease Prevention at **Stanford University**, and two independent teams of scientific analysts, recently attempted to reproduce the findings of 18 research articles. The articles, published in *Nature Genetics* in 2005 and 2006, profiled gene expression from microarray data. Despite the authors' claims that the microarray data set was publicly available, the procedures were not detailed enough to allow for accurate reproduction of the findings for 16 of the 18 articles.[1]

Inability to reproduce findings is not an uncommon problem in modern science. Several other independent studies confirm **Dr. Ioannidis'** findings, including a report by researchers at **Amgen** pharmaceutical company, where only six of the 53 studies they tested were reproducible,[2] and an internal report at **Bayer HealthCare**, where results from published data were irreproducible in two-thirds of their projects.[3] As research becomes more complex and the dependency on detail and accuracy grows, there is a need for more clarity in the publication of methods.

Is the lack of progress in scientific publishing affecting the productivity of science? Data from several recent studies would suggest that this is a possibility. So, inevitably we are faced with the question of what can we do to increase the productivity of science? Is the current problem an example of the way science is performed or the way it is published?

Biomedical Research Budgets at Risk Due to Low Reproducibility

A recent article in the *Journal of the American Medical Association* detailed a large-scale biomedical research budget and spending study by the **Alerion Institute**. The authors of the study found that spending on biomedical research, which had doubled over the last century to an all-time high rate of over $100 billion a year in the U.S. alone, has now begun to decline.

The **Alerion** study found that industry is the largest sponsor of medical research, at 58 percent of the spending, followed by a 33 percent contribution from the federal government. This equates to an approximate $30 billion contributed by the U.S. government each year (from agencies like the **National Institutes of Health** and the **National Science Foundation**), and means that the U.S. spends about six cents of every health care dollar on medical research.

Dr. Hamilton Moses, III, coauthor of the study and chairman of the **Alerion Institute**, said "If we're going to be spending $100 billion a year, we'd better have treatments that work over a long period of time against diseases that are important today and will be more important tomorrow."[4]

Dr. Moses and his team also concluded something rather shocking from their study: while spending on biomedical research has doubled over the past century, approval for new drugs and medical devices has stagnated. Possible causes for the productivity shortcomings in biomedical fields have been linked to the current lack of reproducibility in published work. The implication is an incredible waste of resources and risk to research funding. Drug manufacturers rely heavily on early-stage academic research and can waste millions of dollars on products if the original results are later shown to be unreliable. More, when patients enroll in clinical trials based on conflicting data they may sometimes see no benefit, or worse, suffer harmful side effects.

Unlike pharmaceutical companies, academic researchers rarely conduct experiments in a "blinded" manner. This makes it easier to handpick statistical findings that support a positive result. And, in the quest for jobs and funding (especially in an era of economic malaise), the growing army of scientists need more successful experiments published under their name, not failed ones.

So if everyone wants and needs to reproduce experiments, why are duplicative results becoming so elusive? One reason may be that different labs and different materials can produce variant

continued on page 20