

Committee:	Date:
Culture, Libraries and Heritage Committee	27 January 2014
Subject: Digitisation : progress and policy	Public
Report of: Director of Culture, Heritage and Libraries	For Decision
Summary	
<p>This report provides an overview of the progress made in recent years in the digitisation of collections, particularly at London Metropolitan Archives, and outlines plans for further development in this increasingly important area.</p>	
Recommendation	
<p>It is recommended that the report be noted and the proposals detailed in paragraphs 16 and 17 endorsed.</p>	

Main Report

Background

1. Over the past decade a transformation has taken place within society driven by the increasing importance of the digital world in every aspect of our lives. Notably, information, communication, knowledge and content of every kind that used to be transmitted in writing and in print is moving, or has moved, to digital platforms. Our vast documentary heritage – including books, archives, photographs, maps, pictures, and films – is steadily being digitised to make the constant accessible online. Material which used to be viewable in one physical locality can be seen and used, for business, leisure or research, anywhere in the world.
2. Various business models have emerged for this digitisation activity. Sometimes, charities, publicly-funded bodies or businesses have funded it for public benefit so that the content is freely available; sometimes it has been undertaken on a commercial basis and access is charged for either on a cost-recovery or for-profit basis. There are also hybrid models, where free access is provided to particular communities while others pay.
3. The most successful, high-impact digitisation projects have tended to be those based on big, simple ideas which give access to large coherent bodies of content; examples are given in Appendix 1. Increasingly, these sites become hubs for interaction with their audiences by inviting input, crowdsourcing, blogging, personalisation, etc; they become a kind of social media.

Current position

4. Within CHL, digitisation activity has been centred around the collections at London Metropolitan Archives (LMA), although some digitisation has also been undertaken with the printed holdings of Guildhall Library. LMA has led on three strategic initiatives to digitise its collections and make them available online. They are:
 - *Collage* (graphic collections)
 - Major genealogical collections
 - London collections.

Further information on each of these strands is given below.

5. ***Collage***: LMA holds some several hundred thousand graphic items relating to London, the largest collection of such items held about one place in the world and as such of huge significance for the history of the capital. *Collage* was established in the 1990s originally and primarily for the City's paintings collection but over the past 10 years has been developed into a wide ranging digital image system which holds over 250,000 graphic items from LMA (mainly photographs and prints) as well as the roughly 4,000 Guildhall Art Gallery paintings. *Collage* was a ground breaking service when it was first launched and is now an increasingly known brand name for our image sales and licensing, achieving 2.3 million page (image) views in 2013
6. In 2010 LMA entered into a partnership with the charity New Deal of the Mind (now The Creative Society) which seeks to develop jobs in the cultural sector for young people, and typically those who have been long term unemployed. With funding from the Department of Work and Pensions 26 fixed term placements were set up at LMA to digitise and catalogue some 200,000 photographic images from the archives of the former London County Council and Greater London Council the majority of which are now made available to the public through *Collage*. In 2013 LMA entered a further partnership with New Deal and is currently employing a small number of business trainees from Haringey who are doing further work to enable the uploading of images to the website. The trainees are also digitising the metropolitan print collection, which covers images across greater London and are working on a major overhaul of the *Collage* website which has remained essentially unchanged since its launch.
7. Images created for *Collage* have been successfully repurposed in a number of areas including social media platforms such as Flickr (some 300,000 image views in 2013) and HistoryPin, a site managed by the charity 'We Are What We Do' in partnership with Google to encourage cross generational interaction online. The images have also been reused to generate additional income streams through commercial licensing partnerships which in turn has led to extended services such as high quality print on demand and resale through outlets such as Amazon.com.
8. ***Major genealogical collections***: LMA holds extensive and important collections of archives and printed material used by family historians. These

collections include parish registers from some 1,000 Church of England parishes in London, electoral registers, poll books, Poor Law settlement examinations and wills among many other relevant sources. In 2006, following a report to the then Libraries, Archives and Guildhall Art Gallery Committee, LMA developed a proposal for a commercial partnership to digitise, index and provide an online service to access these archives.

9. After a successful procurement exercise in 2007-2008, a contract was signed between LMA and a commercial partner Ancestry. The new service was launched early in 2009 and to date some 11.5 million images linked to an index of some 77 million names are available for view on the Ancestry website. The service has proved to be immensely popular with users worldwide and in 2012-13 some 22 million views on LMA held material were recorded by Ancestry. LMA is currently working with Ancestry to digitise some more of the family history resources.
10. **London collections:** LMA's remaining collections, that is those which are not graphic and so fall under the scope of *Collage* or major genealogical sources falling under the scope of the Ancestry partnership, are the largest group here and cover over 80km of archives (measured in linear shelving). Not all these collections are owned by the City and collectively are not currently as heavily used as the major genealogical sources.
11. As part of its partnership with Ancestry, it has been possible to have digitised the top level documentation from the London government collections. These include in their entirety the City of London Corporation Court of Aldermen Repertories and Court of Common Council Journals; the main council minutes for the Inner London Education Authority, the Greater London Council, the London and Middlesex County Councils and their predecessors together with any extant indexes. The information in these series is a highly significant and important source for London's history and consists of some 600,000 images dating from 1495-2000. This material is not yet mounted on the web and will make a substantial and valuable addition to the digitised resources available, although a single overarching index is needed to make them properly accessible. A number of other specific collections have been wholly or partly digitised via various academic or other partnerships.
12. It should be added that LMA, like all archive services, also faces the challenge of ingesting, preserving and providing access to 'born digital' records. LMA already manages 70,000 born digital records for depositors, although some material has been received from City teams, and is looking to develop new systems to manage and make accessible its digital archive. This work is at an early stage and further progress reports will be made in due course. Staff are also working with the City of London Police on a business records digitisation project and it is expected that the results of this work will be useful in informing the future development of digital records systems at LMA.

Options

13. It should be clear that the City Corporation is the steward of outstanding, extensive collections that capture the history of London and Londoners from the earliest times to the present day. In addition to the archives, maps, books and images at LMA, Guildhall Library and the Guildhall Art Gallery, there is similar material at the Museum of London who also cover the artefactual side of London's history with every kind of object, reflecting the lives of Londoners since Roman times. If this heritage material is taken in its totality it can be said to represent a sizeable proportion of London's recorded memory.
14. An opportunity therefore exists for a transformative-scale digitisation project to make London's history accessible to the world at large. This totality of material, or a comprehensive subset of it, has the potential to create one of those simple, big idea digital databanks with high impact. This idea has been discussed with the Director of the Museum of London who has indicated a willingness in principle to participate in developing it.
15. We could expect such a resource to be used by millions of people, all round the world. The content would be used for educational purposes, by family historians, academic researchers, designers and artists, architects and planners, and many others. It would also provide a partnership opportunity to other custodians of London historical material including heritage services operated by London's borough councils who could be invited to participate and contribute material.
16. In order to progress this idea, more work needs to be done to provide detailed proposals around funding partners, business models, timescales, prioritisation and selection of content, and technical methodology. Subject to Member approval, the funding to undertake a feasibility study to scope this work will be sought via the Policy & Resources Committee in spring 2014. At this stage, however, it is possible to recommend a number of principles which should underpin any such large-scale City-sponsored digitisation project:
 - Content should, at the basic level, be free to access and use for non-commercial purposes. This does not preclude charging for commercial use (e.g. advertising, publishing reproductions), or supplying high-quality images. The income-generating potential for access to material like this is not great (this has been market-tested with publishers) but the reputational gain and profile-raising associated with opening it up to the world are positive opportunities for the Corporation.
 - Any web/online application pages through which any such digital material is viewed should be clearly branded as originating with the City of London.

Conclusion and recommendations

17. LMA has a proven track record of significant input into the digitisation of its collections. The opportunity now exists to look at the creation of a larger scale project which would be not only philanthropic but also forward-looking, making a statement about our commitment to a digital future, and to Digital Britain (a stated government objective). Members are therefore asked to note the achievements to date, agree the proposed principles, and support the request

for the funding of a feasibility study to scope the project in more detail, as outlined in paragraph 16.

Contact

Geoff Pick

Director London Metropolitan Archives

Culture, Heritage and Libraries

020 7332 3833

geoff.pick@cityoflondon.gov.uk

David Pearson

Director, Culture, Heritage and Libraries

020 7332 1850

david.pearson@cityoflondon.gov.uk

Appendix 1

Examples of large-scale high-impact digitisation projects

Google Books: <http://books.google.com/>

- Aspires to create digital surrogates of all the world's books, by encouraging authors/publishers to submit content, and through a huge retro-digitisation programme in major libraries in the USA and Europe
- Freely available to use, funded by Google in pursuit of their mission "to organise the world's information and make it universally accessible and useful"

JSTOR: <http://about.jstor.org/>

- Creates digital versions of back runs of thousands of journals, extensively used by libraries and also individuals worldwide
- Operates as a not-for-profit organisation, and charges subscriptions for access based on cost recovery and sustaining the operation

National Library of Norway: <http://www.nb.no/English/The-Digital-Library/Digitizing-policy>

- Is digitising the entire National Library of Norway collections
- Will be freely available to use, funded by the National Library in pursuit of its vision "to act as a national memory bank by providing a multimedia knowledge centre ... a modern national digital library"

Census records: <http://www.nationalarchives.gov.uk/records/census-records.htm>

- Searchable database of all the census records for England and Wales, 1841-1911
- Managed by The National Archives in partnership with several commercial digital publishers. Websites can be freely searched, but non-subscribers typically pay for downloading or obtaining full transcriptions of documents

British Newspaper Archive: <http://www.britishnewspaperarchive.co.uk/>

- Partnership between the British Library and the publisher brightsolid to digitise 40 million 19/20th c newspaper pages
- Free to access within the BL reading rooms, otherwise charged; some content will be free to access from HE institutions because JISC provided some funding

Old Bailey Online: <http://www.oldbaileyonline.org/index.jsp>

- Fully searchable, digitised version of proceedings of all Old Bailey trials 1674-1913
- Free to access; funded from a series of grants from the Big Lottery Fund, AHRC, ESRC and JISC, led by a consortium of university-based academics

Early English Books Online: <http://eebo.chadwyck.com/home>

- Full-text digital surrogates of every book printed in Britain before 1701
- A commercially-created project; extensively used in academic/research libraries but access depends on subscription

Genealogical records: <http://www.ancestry.co.uk/>

- The biggest of a number of genealogical databases, which aggregates records from a wide variety of sources, internationally
- Access is by subscription, by libraries or individuals; the LMA records on Ancestry are freely searchable from LMA and other CoL sites