Bridging the Gap

Investigating Community-Led Image Collections

Project Report

June 2006

Jonathan Miller & Peter Robinson,
Learning Technologies Group, University of Oxford

Rupert Shepherd,
Ashmolean Museum, University of Oxford

Karla Youngs & Grant Young,
Technical Advisory Service for Images

Contents

1 CLIC summary

The Community-Led Image Collection (CLIC) scoping study was commissioned by the JISC to review current community image collection activity. It was to make recommendations about how national initiatives could help embed collections within the wider educational community, to provide better access to high-quality educational image material.

The study, which was a collaboration between the Learning Technologies Group, at the University of Oxford, and the Technical Advisory Service for Images (TASI), at the University of Bristol, undertook the following activities:

- An overview of current community image collection activity, by consultation with collection providers.
- Selecting community image collections to use as case studies.
- Surveying image owners and users via workshops and surveys.
- Investigation of viable models for a network of image collections to promote sharing of image resources.

1.1 Outcomes

The outcomes of this study are a model for a network of community-led image collections, and recommendations that should be implemented to promote sharing of image material.

We propose that image collections, based on subject area, be established for the exchange of digital images within the wider educational community. Images will be shared in a devolved, self-organising manner, by communities of practice (Lave and Wenger), under Creative Commons licences (or similar) with initial material provided by current community-based image providers and owners that are amenable to this model. This initial material must be clearly identifiable in terms of ownership, copyright and licensing.

These community collections will integrate with local collections and national collections in a three-tier network, sharing catalogue information, and allowing cooperative classification of material held in other collections.

We believe that a social support network, which can disseminate information, and offer advice at institutional, subject, and national levels must underpin these collections.

The three-tier model may not suit all situations due to risk factors associated with different image material and it is proposed that further work be done at a national level to create policy for sensitive material such as clinical medical recordings (See the CHERRI-PIE report).

1.2 Key findings

1. Local needs are best met locally, with support from subject-based communities of practice to provide core copyright-cleared material.

   National initiatives should be driven by subject-based demand and dissemination.

Community-owned repositories, containing material in their own right, if linked through a directory space, will also provide subject-based single points of access to relevant material in local and national collections.

2. A register of visual teaching material needs and requirements should be identified as soon as possible. Bodies that are recognised nationally as subject specialists (e.g. Higher Education Academies, Arts and Humanities Data Service) should work to compile the register, and build up small corpuses of core subject teaching material to satisfy these needs.
3. Social communities of practice need to be identified and nurtured until they are mature. Ideally this would be done through subject specialists who are well qualified and best placed to undertake the task. Nominated liaison contacts and mailing lists need to be set up and coordinated by the Higher Education Academies (HEA), the Resource Discovery Network (RDN), the Technical Advisory Service for Images (TASI) and national initiatives such as the Arts and Humanities Data Service (AHDS).
4. Creative Commons licences, or a limited number of Creative Commons-style licences, should be adopted as the default licensing system for born-digital, nationally funded initiatives. These must always accompany the images. This will alleviate the major hurdle of confusion over Intellectual Property Rights (IPR) that prevents the sharing and modification of material.
5. Material that is deemed sensitive, either for ethical reasons (such as Clinical Recordings in Academic Non-Clinical Settings, photographs of children), or disputed IPR, or data protection and privacy issues, should be contained within local repositories under local authentication and localised access control, until and unless a national system is devised.
6. A series of community-owned directories of image collections, in which collections would register themselves, should be set up immediately as a first step towards coordinating future cooperation within communities. After initial seeding, these should be maintained and supported by the communities, with minimal system administration. However, communities must feel that they, and not national institutions, are the owners of these directories.
7. Partnerships for sharing subject material must be explored across the education, museum and commercial sectors at national and international levels. This would be aided by institutions releasing material under Creative Commons-style licences.
8. High quality metadata is vital. This contextual information can only come from the communities themselves. Mechanisms for post-submission annotation and update by communities will prevent collections from becoming static archives. Systems need to be built into repositories that allow academic peer reviewed social cataloguing.
9. The CLIC three-tiered model being proposed assumes that there is some interoperability between levels and across subjects, via metadata cross-searching or aggregation. This would be a challenge to implement within a distributed, community-owned structure and we did not find underlying infrastructure mechanisms in existing community image collections. While efforts should be made to ensure compatibility, it may prove more efficient to concentrate on more lightweight discovery mechanisms such as Really Simple Syndication (RSS) or the Resource Description Framework (RDF) that are already in widespread use and could readily be adopted by community image collections. This should be seen as a first step towards implementing

more sophisticated systems such as the Open Archives Initative's Protocol for Metadata Harvesting (OAI-PMH).

10. A national collection of material does have some support from image collection owners and users but should be initiated with minimal staffing, aiming for devolved community input and support along subject lines. Any national collection should seek to incorporate material from community collections in the first instance, after which its role would be to supply images that are not available elsewhere, under suitable licence terms.

1.3 Proposed three-tier model

During the study we found that the biggest barrier to sharing images among individuals, institutions and communities is lack of trust. Technical barriers can be overcome, but a technical solution in no way guarantees the success of image sharing. The problem is social, not technical. Collection providers need to be able to trust that the material they provide will only be used in ways that they find acceptable, and that their ownership and IPR will be respected.

Trust needs to be built within and between institutions, so that image-based teaching materials provided by one party will be made available for use by another, and that all parties can rely on the veracity of information provided in image collections.

Trust is particularly important when dealing with high-risk material such as images of children or those that can be used to identify individuals. Images that are subject to copyright should be clearly distinguishable from those in the public domain and any restrictions on the use of images should be easily understood. This will allow collections to judge the risk involved in accepting content, or information from another source, and presenting it along with their own material. The risks associated with different types of material vary dramatically. National collections do not want sensitive high-risk material, such as clinical records, alongside their own low-risk material. Conversely, local repositories have the responsibility not to disseminate any high-risk material they may hold.

In order to make low-risk material available, while at the same time restricting access to high-risk material, our study proposed the establishment of community-based image collections that are situated between existing local and national collections, creating a three-tier system:

1 Local collections of high-risk material held within trusted communities of practice;

2 Open-access community repositories with devolved management of low-risk submissions from HE/FE. These communities of practice will only hold material that is rights-cleared under Creative Commons (or similar) licences. Management will be devolved to self-organising subject-based communities of practice with minimal staff support;

3 National Image Collections incorporating rights-cleared low-risk material from the community repositories, and reciprocating by providing low-risk, rights-cleared material from their own holdings.

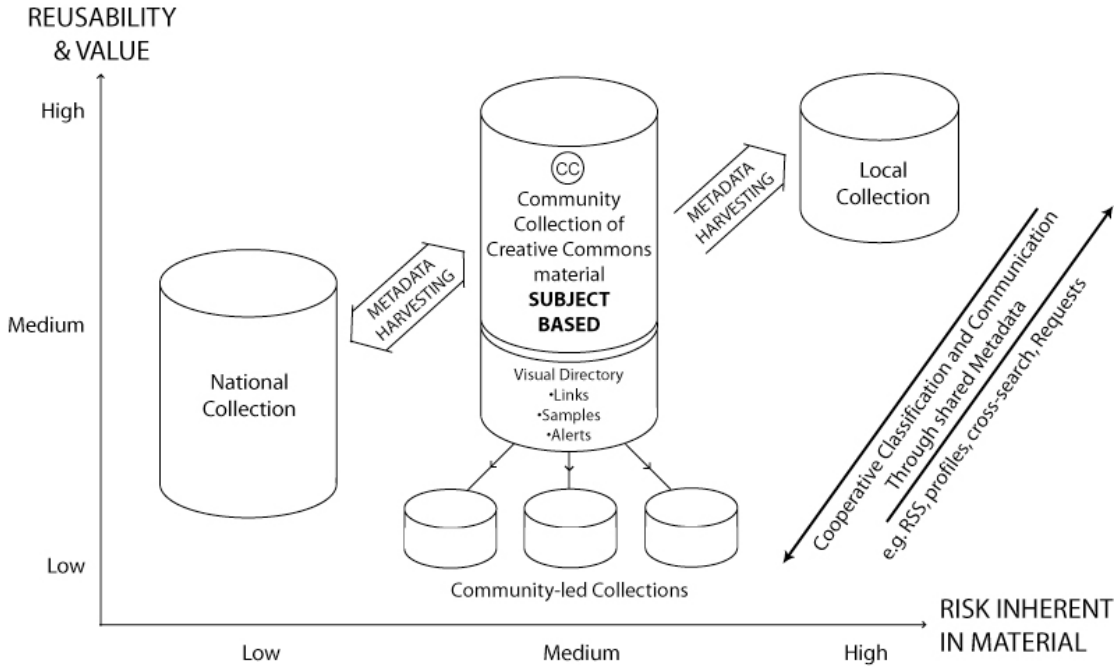

Figure 1 Proposed three-tier model for linking image collections

In order to function together, all three systems will require the common attributes of:

- a system for the syndication and aggregation of discovery information;
- local authentication and authorisation controls for high-risk material;
- metadata harvesting mechanisms;
- consolidation of marketing opportunities;
- community-led requests for additional material;
- cooperative classification and communication through shared metadata;
- higher-risk material to be held locally, lower-risk material should be shared;
- implementation based on open standards.

1.4 Recommendations to provide a social network foundation

1. The establishment of an emailing list containing an image liaison contact per institution.
2. Each Higher Education Academy subject network should nominate an image liaison contact.
3. The establishment and publicising of a directory of image collections, to which additions can be made by self-registration.
4. The establishment of an annual conference, forum or meeting, focusing on image collection needs.

1.5 Recommendations for the discovery of image material

1. Any proposed network should increase opportunities for communities of practice to communicate their needs for image material to the relevant funding bodies.

2. Any proposed network should offer opportunities for members of communities of practice to find material they need and discuss the material with fellow users.
3. Guidelines should be issued at national and institutional level, giving clarification of appropriate licences governing the reuse of image material.
4. The RDN subject centres should expose their catalogue information on image collections prominently on their websites.
5. The HEA websites should link to both the proposed central directory and the RDN image collections listings.

1.6 Recommendations related to rights

1. All JISC-funded repositories should offer Creative Commons licence (or similar) options on upload of material.
2. Search for CC-style material should be offered on all nationally-funded repository collections.
3. Guidelines for the use of CC-style licences should be issued for publicly funded image digitisation projects.
4. Individuals in institutions should be given clear guidelines on the IPR status of their work.

1.7 Recommendations related to technical needs

1. Designated subject user communities should have funded ability to keep track of their wants and technical requirements for national-level solutions to the costly areas of interoperability, and preservation and archiving of images and their metadata.
2. JISC should work to bridge the gaps, identified by recent reports, in the workflow for creation and use of images in educational contexts. Key areas for further development include metadata interchange, archiving, Search/Retrieve via the Web (SRW) and the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), persistent identifiers and authentication and authorisation to national schemes (perhaps utilising protocols such as Shibboleth).
3. JISC should fund the creation and maintenance of software libraries, toolkits and modules that address the needs identified above, to enable small image collections to interact with the JISC Information Environment and the proposed three-tier model. These toolkits and modules should be incorporated into existing open-source image software to make them compatible with the JISC Information Environment and the proposed three-tier model. (See section 7.3)
4. Provide comprehensive training and technical support for image collection providers. This should be achieved through existing bodies such as TASI, the Centre For Educational Technology Interoperability Standards (CETIS) and UKOLN, and should use the standard toolkits developed above.
5. Grant awarding bodies should be encouraged to retain the scripts and technical infrastructure behind the digitisation projects they fund, as a project output that should be made available under Open-Source licences.

1.8 Recommendations related to risk

1. Guidelines should be issued to help collection owners understand the risks inherent in different types of material and the responsibilities associated with storing image material in publicly accessible collections.
2. Investigate further the CHERRI C+L Model (see Chapter 8) for clinical recordings or other high-risk material that is bound to a community.

2 Current landscape and context

Repositories and collections of digital images have flourished in recent years, supported by national archive initiatives, research funding bodies' requirements for automatic deposits and considerable National Heritage funding. At the same time, advances in affordable digital photography and home scanning have empowered teachers and researchers to create their own digital images more easily and store them, in an ad hoc manner, on their local machines. However, usage of image collections in teaching and learning has not, perhaps, flourished to the same extent as the rise of digital photography. A number of technical initiatives have been funded to provide common gateways or portals to nationally hosted image collections and there is concern that these are still not being used as widely across the educational community as was intended.

The aim of the Community-Led Image Collection (CLIC) scoping study was to suggest how to provide better access to high-quality image material that is relevant to the needs of the educational community.

2.1 Impact – on teaching – of networked collections

Presentation of the visual material available in existing slide collections and personal digital image collections forms a central part of the teaching programme in a wide range of subjects. Digitisation, networked management and distribution of visual resources have the potential to transform their use in teaching, offering:

- efficiencies in construction of lecture material;
- flexibility in selection of material in response to discussion in seminars and classes;
- full availability of lecture material to students for follow-up study;
- enhancement and customisation of visual material.

An online repository of digital images and contextual information that could be downloaded to a local system would allow a range of new teaching activities to be realised. Additional functionality provided in many systems would allow distributed activities impossible with analogue material. Potential teaching activities could include:

- user annotation of images
- web-based slideshows
- comparison of two or more images on a display
- packaged slideshows for offline presentation
- printable flashcards of images and accompanying notes
- incorporation of remote collections
- uploading of personal images
- search and browse function and cross-collection searching
- generation of personal collections by academic users
- custom catalogue data structures
- slideshow light tables.

2.2 The 'discovery gap'

There are obvious parallels between any information-seeking activities such as searching for journal articles for research or searching for image material for a teaching project. The user wishes to find material quickly, to ascertain the validity of the source and to understand how they may make use of the material.

There is currently a gap between how users search for information, typically through Google, and the provision of high-quality academic resources and national initiatives by local support services such as their library or virtual learning environment, and subject-based organisations such as the HEAs or RDNs. Often, resources are locked away inside protected systems that prevent discovery by those outside the target community, and especially commercial search engines.

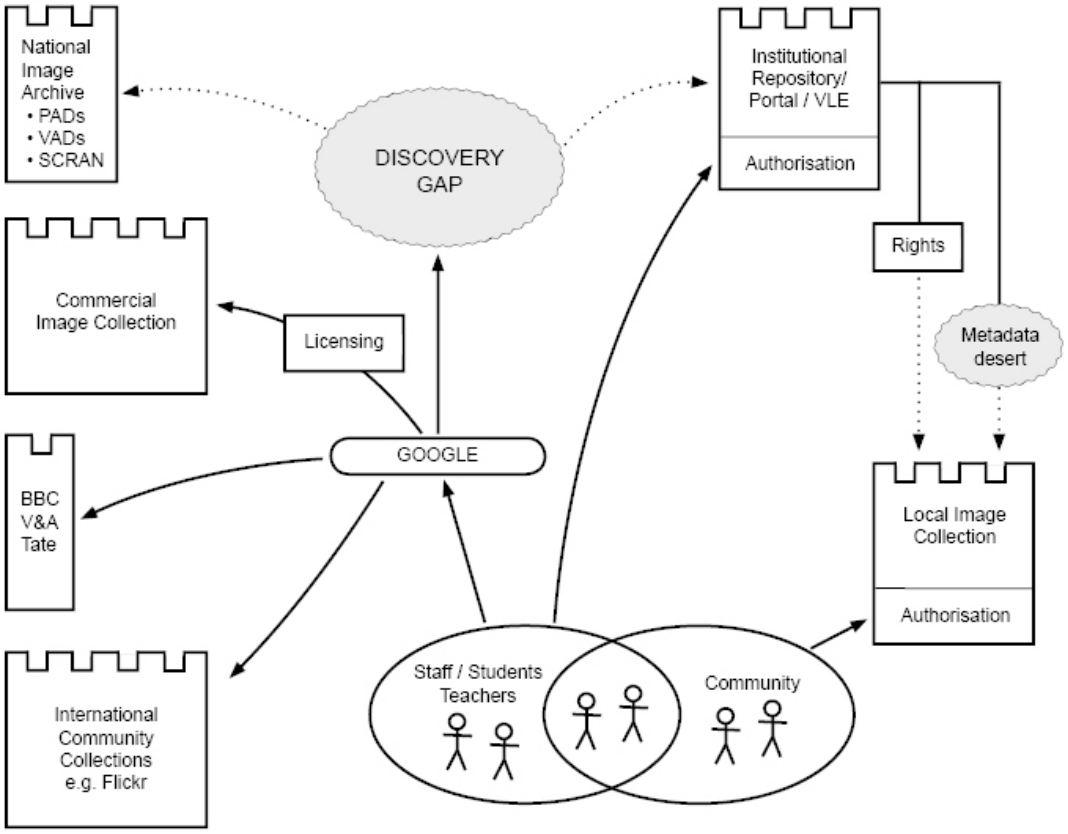

Figure 2 The discovery gap. Which route shall I take to find materials?

Figure 2 shows various sources of image material available to the academic community. An academic who wants to find image material may use Google or refer to institutional or local collections. If they cannot find the material they want, they very quickly give up. Barriers such as authorisation (from library membership, to online subscription) prevent access by both search engines and 'unauthorised' users. Internet search engines serve the commercial sector, and are not geared towards providing high-quality educational resources, so many of the collections funded recently do not feature prominently in their search results. This leads to what we describe as the discovery gap – the educational community cannot easily locate images that have been provided for its use.

Very few of the image resources that reside in local community collections will be discovered by users from outside that community. This is often because institutions

and larger collections are wary of incorporating material that may not meet their own standards in terms of metadata, image quality, rights clearance. In Figure 2 this is shown as the ‘metadata desert’.

2.3 Increasing access

The majority of image users start their search for images with Google's image search (http://www.google.co.uk/imghp). This service, and others like it, allows searchers to find pictures quickly and easily by entering keywords into a web page. Images are presented as thumbnails, and a link is provided to the hosting site (along with a reminder about potential copyright issues). The images, by and large, do not have sophisticated metadata accompanying them, but the popularity of Google's image search suggests that it satisfies the majority of people the majority of the time.

Flickr (http://www.flickr.com/) is one of a number of community image sharing sites. It allows registered users to upload images that they own onto the web so that they can be shared with friends, family and colleagues. It allows images to be grouped, and tagged using simple keywords. Again, the metadata that accompanies images is not sophisticated, and often it is inaccurate. Flickr is a community-led site that is innovative in that it promotes the use of Creative Commons licences and encourages both image owners and viewers to annotate images. By satisfying the needs of their audience these image sharing sites have driven up the access to, and use of, image material on the internet.

Educational collections typically see their purpose as being to provide high quality resources that are well documented and authoritative. They rarely incorporate the features that make Google and Flickr popular, but they could learn many lessons from these sites. In order for material to be used, it first has to be discovered. By limiting access to collections by password, excluding search engines, preventing user annotations and failing to define appropriate terms for acceptable reuse of material, collections that are aimed at the education sector are reducing their value to their target audience.

2.3.1 Creative Commons

Creative Commons (CC) is a non-profit making organisation devoted to expanding the range of creative work available for others legally to build upon and share. The Creative Commons website (http://www.creativecommons.org/) enables copyright holders of images and other material to grant some of their rights to the public, while retaining others, through a variety of licensing and contract schemes including dedication to the public domain.

The intention is to avoid the problems current copyright laws create for the sharing of resources such as music and images. The project provides several free licences that copyright holders can use when releasing their works. Creative Commons also encourages the use of RDF/XML metadata that describes the licence and the work, making it easier to process and locate licensed works automatically. The licences have been successfully localised for over 26 countries around the world. A particular strength of this CC licence work lies in the provision of supporting icons and labelling materials that teach the general public to recognise CC material.

The following list, taken from the CC website, describes each of the six main licences offered when you choose to publish your work with a Creative Commons licence. The licence types are listed starting with the most restrictive and finishing with the most accommodating. Next to the title of the licence is the number of images held in the Flickr social image sharing site as of 3 March 2006. This statistic perhaps reflects which licences are most commonly chosen by the public.

Attribution Non-commercial No Derivatives (by-nc-nd) 2,588,115 images in Flickr	This licence is often called the 'free advertising' licence because it allows others to download your works and share them with others, as long as they mention you and link back to you, but they can't change them in any way or use them commercially.
Attribution Non-commercial Share Alike (by-nc-sa) 2 580 522 images in Flickr	This licence lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.
Attribution Non-commercial (by-nc) 1 170 629 images in Flickr	This licence lets others remix, tweak, and build upon your work non-commercially. Their new works must also acknowledge you and be non-commercial.
Attribution No Derivatives (by-nd) 250 487 images in Flickr	This licence allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.
Attribution Share Alike (by-sa) 645 582 images in Flickr	This licence lets others remix, tweak, and build upon your work even for commercial reasons, as long as they credit you and license their new creations under identical terms.
Attribution (by) 869 989 images in Flickr	This licence lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation.

2.3.2 The Creative Archive

The Creative Archive is a collaboration between the BBC, the British Film Institute, Channel 4 and the Open University, to produce a system for releasing their archive material to the public. The Creative Archive Licence is very similar to the Creative Commons licences, with some additional restrictions:

- UK: the Creative Archive licence is only for use within the UK;
- Endorsement is forbidden: you must not use the work and/or derivative work in any way that would suggest or imply the licensor's support, association or approval;
- Attribution: you must give credit to the author of the work;
- Share-alike: you must license derivative works under the same licence terms;
- Logo: you must attach the Creative Archive logo to all derivative works;
- Non-commercial: you may make personal use or use for educational purposes within any educational establishment [listed elsewhere], but may not make any commercial use (including professional, political or promotional uses);

The Creative Archive Group publish download figures for their licensed material, from which it appears that the scheme is becoming increasingly successful. There is evidence of the licence becoming used more widely by Government-sponsored initiatives; for instance the free-to-view, UK Government sponsored Teachers TV channel (http://teachers.tv/helpCopyright.do) hosts material from shows online under a provisional Creative Archive licence, thus enabling teachers to reuse material without worry.

It is worth noting that the 'educational use' clause in their licence is restricted to institutions that are in a set list of UK schools colleges and universities. This reflects the difficulty of defining 'education', especially in this age of life-long learners.

2.4 Increasing relevance

It is generally considered that providing high-quality material is not enough to guarantee that resources are used. They must be perceived as meeting the needs of users. A recent review of digital repositories work (Heery and Anderson, 2005) concluded:

> Repository interfaces should be directed towards 'communities of practice' and more effort should be made to tailor services for specific user communities, rather than producing generic interfaces. This is likely to be particularly important to encourage take-up within FE, although it is still important with HE … Repository developments should, depending upon their primary focus, relate to the processes and practices of research, teaching or learning – buy-in from the community is unlikely to be extensive unless this happens. JISC should identify current practice of researchers, teachers and learners, and seek to base services on supporting their needs.

It has also been suggested by the JISC Exchange for Learning Programme (X4L) manager, investigating re-purposing and sharing of existing and forthcoming content, that teaching and learning repositories need to focus on delivering to communities of practice, if significant take-up and use of the content is to be achieved.

2.4.1 JISC Information Environment

On a national level, the JISC Information Environment (http://www.jisc.ac.uk/index.cfm?name=ie_home) has emerged in response to the challenge of supporting and coordinating the variety of nationally-funded information initiatives from a technical and social perspective. It will provide a range of services, tools and mechanisms for colleges to exploit fully the value of online resources and services.

Key strands in the Information Environment programme include:

- accessing online resources through portal services
- building and sharing community resources
- developing shared services
- promoting good practice in the preservation of digital resources

2.4.2 The JORUM national repository

JORUM (http://www.jorum.ac.uk/) is a free online repository service, supported by a range of national organisations, funded through JISC for teaching and support staff in UK Further and Higher Education Institutions. It is helping to build a community for the sharing, reuse and repurposing of learning and teaching materials. Although currently (January 2006) it has few deposited materials and is not yet offering a full service, the repository is likely to receive illustrative material. The repository has done a lot of work on the licence policies necessary for material uploaded from

institutions, including some study of the applicability and difficulty of using Creative Commons material in an institutional setting. The repository does not appear to have a specific type of object that it wishes to collect:

> Contributors can make their materials available to the community easily and freely. Materials are stored and catalogued in the system, making them accessible to a national audience of Further and Higher Education teaching staff. Materials can be downloaded and repurposed by Users, who are also able to provide comments and feedback on deposited work.

JORUM requires only minimal metadata submission by the contributor, but requires a statement of IPR. A cadre of cataloguers then completes the catalogue record, which is finally sent for peer review and possible amendment.

2.4.3 The CREE project

The Contextual Resource Evaluation Environment (CREE) project, based at the University of Hull and the University of Oxford, studied user preferences in searching for information, focusing on library-based services and portal interfaces. CREE investigated, through large-scale user surveys and focus groups, how users wish to see the services they requested presented to them in portal and non-portal environments such as virtual leaning environments (VLE). It has made recommendations for how national search interfaces might be improved to meet local needs; their findings will be incorporated into the CLIC project.

2.4.4 The Digital Picture

The recent project, the Digital Picture, a UK-wide initiative to explore digital image issues in the visual arts education community (http://www.thedigitalpicture.ac.uk/, Arts and Humanities Data Service, 2005), surveyed user needs across the sector. Although the specific issue of community involvement was not raised in the questionnaire, several respondents – presumably reflecting the opinions of their colleagues – stressed how vital it was in their general responses to the consultation. One noted that they needed “Support for heterogeneous networks that allow teachers and researchers access to the hardware and software tools that are most suited to our needs, not those of the IT managers”, whilst another put the issue particularly clearly: “I’m pleased by your demand-driven, rather than supply-driven approach, since other initiatives don’t seem to have learnt from the UK e-University mistakes.”

2.4.5 Keeping up with technology

The cost of entry to digital image creation has fallen. The rise of easy-to-use, high-quality consumer capture devices, such as scanners and digital cameras, allows educationalists to create their own visual teaching material without recourse to specialist services. Desktop computer software that enables non-technical users to manage and keep track of their digital images is in widespread use, and services that allow the upload and sharing of personal image material are growing in popularity.

Any system that seeks to incorporate home-grown ‘born digital’ material must incorporate these low-end systems into its workflow; it must address the current needs and practices of the community that it serves.

It is interesting to note, alongside Kodak's recent decision to stop making slide projectors, that Nikon recently announced that it intends to concentrate on the digital picture market and will be ceasing production of the majority of its film camera range. This reflects an increasing acceptance of digital technology, not just at the consumer level but also by professional photographers. There is also currently a huge increase

in social image file-sharing forums, such as flickr.com, which allow the public to share images with family and friends and with a wider audience. It is in this area of social file-sharing of images that most of the newer techniques for accessing, networking and cataloguing images and their metadata are taking place.

The education sector would do well to attempt to incorporate techniques that are commonly employed in the commercial image-sharing sector to build thriving vibrant communities.

2.5 Overview of current community image collections

This CLIC scoping study surveyed the growth of community-led image collections by contacting image collection providers. The study investigated the socio-cultural and technical barriers these image providers have faced in building collections.

It is important to understand the technical solutions to which authors have turned at a local level and to see what obstacles they faced. Can their needs be met by national initiatives? There is potential for material from these communities of practice to be opened up and shared more widely, both within subject areas and at a national level. For this to happen, the material must be:

- eligible for general use;
- easily discovered through a common interface such as a web-browser;
- subject to agreed, simple-to-understand terms from the licence holders.

2.5.1 A matrix of community-led image collections

CLIC compiled a list of nearly 500 UK image collections, using information provided by TASI's Image Collection Registry, the Resource Discovery Network (RDN) and various funding bodies. These were mainly arts and humanities collections, perhaps because the science communities have already created large community portals to meet their needs, in which images are stored as part of general research (see EMBL). The majority of the collections that we found were not 'image databases'; rather, they were websites that contained significant amounts of information that put the image into context.

It was decided at an early stage to restrict the scope to UK-based collections. This decision was made for two reasons. First, from a practical point of view, there were simply too many collections worldwide to log each one. Second, non-commercial collections based outside the UK are problematic due to the differences in copyright law, licensing terms etc.

Funding bodies provided an initial assessment of the range of projects that have been developed by communities, for communities. The following organisations were used during this phase:

- Arts and Humanities Data Service (AHDS)
- Enrich UK (lottery funding)
- Arts and Humanities Resource Council (AHRC)
- Museums, Libraries and Archives Council (MLA)
- Resource Discovery Network (RDN)
- Higher Education Academy (HEA).

The TASI Image Collections Registry (ICR) contains details of several hundred image collections and, together with collections flagged by funding bodies, these formed the backbone of the research.

It is important to realise that these collections almost certainly hold only a small fraction of the material that is held in inaccessible and undiscoverable locations such as within institutional and departmental servers, Virtual Learning Environments, repositories and material held on desktop PCs. In 2003, research at the Penn University by the Visual Image Users Study (http://www.libraries.psu.edu/vius/) estimated that most of the institution's image assets were held on the desktop, and were completely inaccessible by the wider university.

Following discussions between the project partners on the objectives of this work, it was agreed that the matrix would record the following information:

Contact details

Collaboration

Organisations that funded the project were recorded, as were any organisations working in partnership with the project leaders.

Subject scope

This field was used to record any HE/FE subjects that would explicitly benefit from the image collection. This provided an overview of the level of support currently available for each of the subject areas.

Subject keywords

Keywords were recorded to provide more specific information about the content and nature of the resources

Collection size

The number of images currently available for viewing online was recorded.

Timescale

The start and finish dates of the projects were recorded to give an indication of the current and future activities of each project

Intellectual property

Guidelines on the permissible use of the images and the copyright ownership were recorded. This information was drawn solely from the copyright notices of the institution.

Restrictions

Information was recorded about any restrictions based on financial requirements, registration requirements and restrictions based on an individual’s profession. A separate field clarified the nature of these restrictions.

A free-text description field captured all remaining information

The matrix of collections is presented in Appendix 11.2.

3 Current collections

3.1 Image collection provider survey

In liaison with TASI, and after consulting with stakeholders and case-study projects, we constructed a survey of 49 themed questions designed to be delivered to as many as possible of the 500 community-based image projects identified in the matrix of collections (WP2). The survey investigated the attitudes of image collection providers to the obstacles and hurdles that they face in building collections, and any needs that could be met by national initiatives. The survey also posed questions about the access and rights that govern the use of their material, and the attitudes of collection owners to sharing their material with educational users. The survey was completed in January 2006 and 81 image providers completed the online and paper-based survey. The survey gathered the following data:

- the intended audience of the repository, and the actual audience
- the rate of growth of the repository, its perceived lifetime, and the size of its audience
- community-nurturing features such as commenting, forums, user additions
- access restrictions and rights management
- content monitoring and content quality
- obstacles faced during setup, and the technical solutions adopted
- needs that could be met by national initiatives
- mapping of attitudes to sharing material with Creative Commons licences.

The full survey took place from October 2005 to January 2006. See Appendix 12.4 for an example of the survey questionnaire and Appendix 11.5 for the full results, with expanded discussion.

Research questions posed by the survey

- What are the key barriers to growth, and what are the technical support needs that could be addressed centrally?
- What missing technical infrastructure needs could be solved by a national CLIC network or service?
- What metadata exists locally that could be exposed to a national service (portal or directory)?
- Is the collection part of a community? Is there a demand for sharing material?
- What are the collection owners' attitudes to sharing and allowing open access for educational use?

What are the key barriers to growth, and what are the technical support needs that could be addressed centrally?

The major hurdle in collection building was lack of time, cited by 56% of the providers; next was insufficient funding at 54%, and then lack of technical knowledge. When asked what services could be provided nationally to help, the most commonly identified need was technical support at 46%, followed by marketing at 42% and then search facilities at 41%. The latter two are interesting as they again suggest support for a directory of collections held nationally to promote discovery, or for a portal that allows cross-searching. In further comments, the providers mentioned,

unsurprisingly, that lack of funding and the difficulty of raising funds after an initial digitisation project had ended were a barrier to growth.

What missing technical infrastructure needs could be solved by a national CLIC network or service?

In the survey, the image providers were asked to look at three models that could be instigated at a national level: a directory of collections, a directory of catalogue information that is held centrally (perhaps through a portal), and a national collection of image material. Respondents to our survey were wary about how a potential centralised service would affect them, expressing the fear that a centralised system would lead to marginalisation of smaller collections. The responses to our question about a centralised service showed that:

- 84% support a central directory of image collections
- 76% support a service holding metadata from image collections
- 31% support donation of images to a central image repository/collection.

What metadata exists locally that could be exposed to a national service (portal or directory)?

On the whole, the image providers did not have sophisticated systems that would allow metadata to be harvested or exported easily. This means that, currently, the exposure of collections' catalogue information for cross-searching would be difficult. The metadata standard most commonly mentioned by providers was Dublin Core, within the context that their own internal cataloguing could be mapped to this system.

Is the collection part of a community? Is there a demand for sharing material?

Nearly all the collections surveyed could be said to be part of a community, ranging from a narrow subject area or research topic to wider subject areas of interest to the general public. The image collections comprised the full range of disciplines covered by the JACS subject coding. There was considerable coverage of historical and political subject areas, due to the large number of museum and heritage image collections that contained digitised historical material. *Creative arts and design* was the second most commonly covered subject area. The areas of *mathematics* and *business and finance* had the least image provider coverage. It should be noted that the image collection providers often selected multiple subject areas to which they felt their image collection was of value. One particularly large collection covering a broad spectrum relating to *medicine* was felt by the provider to be of use to a number of subjects, including *history* and *design and technology*.

The majority of the providers accepted donations from people with material relevant to the collection, with some actively seeking out material. Around 62% had contributions from people outside the immediate organisation. Surprisingly few collections allowed viewers to annotate the material online.

What are the collection owners' attitudes to sharing and allowing open access for educational use?

This was perhaps the trickiest question to resolve because of the wide range of reasons for placing material online, visions for the collection's long-term future and of definitions of educational use. The vast majority of collections surveyed were open-access web collections that allowed an unregistered guest to view the thumbnail image and the associated higher-quality image. Broadly speaking, the majority of collection owners were happy for the images to be used for individual research, or by

an individual for educational purposes. However, most respondents felt that sharing images more widely would result in unauthorised commercial reuse, publication without recompense and loss of copyright control. The most common concerns raised are listed in this table.

Common concerns about image sharing	Number of respondents
Loss of copyright control	25
Unauthorised commercial reuse	21
Lack of attribution	8
Ethical concerns	7
Unconcerned	5
Don't own copyright	5

For some providers (5 respondents) there were very few concerns, adopting the attitude that the purpose of the site was to provide greater access to the material:

> My primary concern is to make the images available to as many people as possible.

A common policy was to restrict images on the web to low-quality or small pixel dimensions as a deterrent against misuse. Watermarks themselves were not considered to be popular, the majority had neither textual marks on the images nor formal encrypted watermarks. Only one respondent actually mentioned a formal licence for reuse attached to their material, a Creative Commons (http://creativecommons.org) licence:

> That users adhere to the terms of use a Creative Commons Attribution-NonCommercial-NoDerivs 2.5 Licence.

This concern about misuse was tempered by widespread agreement among most providers that the material could be used for private research or an by an individual for educational purposes – uses already permitted under 'fair dealing'.

3.1.1 Typical community-led image collection

According to the respondents to our survey, a typical image collection could be described in the following manner:

- Built as a bespoke system running on a dedicated server in-house, providing open access via the web.
- Contains over 1000 images and expects to grow to contain 10 000 images within two years.
- Not targeted specifically at education, though it has a policy on the educational use of images.
- Provides information about items in its collection as Dublin Core metadata, if it provides metadata at all.
- Would expect attribution of the collection as the source of any material reused.
- Wants to maintain copyright over the images in its collection.
- Would expect payment for high quality printing or commercial reuse of its images.
- Does not watermark its images, though it states its ownership of the images on the site.

- Considers time, money, and technical knowledge to be the greatest barriers to future development of the collection.
- Would like technical support and marketing assistance from a centralised service.

3.1.2 Coverage

The following table illustrates the subjects that individual image collections felt were related to their material. Note that more than one subject area could be nominated and hence the percentage response does not total 100%.

Q8 Subjects covered by your collection	**Response per cent**
Historical, Archaeology and Philosophical studies	**62.9%**
Creative Arts, Music, Cinema, Photography and Design	55.7%
Social studies, Economics, Politics, Anthropology, Human and Social Geography	51.4%
Architecture, Building and Planning	37.1%
Education	37.1%
Biological Sciences, Biology, Botany, Zoology and Psychology	21.4%
Linguistics, Classics, English studies and related subjects	18.6%
Medicine and Dentistry	17.1%
Nursing, Anatomy, Physiology and Pathology	17.1%
European Languages, Literature and related subjects	15.7%
Eastern, Asiatic, African, American and Australasian Languages and Literature	14.3%
Technologies, Metallurgy, Ceramics and Materials Technology	11.4%
Mass Communications and Documentation, Media Studies, Publishing and Journalism	11.4%
Veterinary Sciences, Agriculture, Forestry	10%
Physical Sciences, Chemistry, Materials Science, Physics and Geology	10%
Engineering	10%
Law	7.1%
Business and Administrative studies, Finance, Accounting, Marketing & Accounting	5.7%
Mathematics, Operational Research, Statistics and Computer Sciences	4.3%

It is estimated that the image collections that responded to the survey included over 1.5 million images. We asked respondents to specify which of the JACS subject categories they considered were applicable to their collections. There is considerable coverage of historical and political subject areas, perhaps due to the large number of museum and heritage image collections that contain digitised material. The 'Mathematical' and 'Business and Finance' subject areas had the least coverage. It is to be noted that the image collection providers often selected multiple subject areas to which they felt their image collection was of value.

One particularly large collection felt it had appeal across a broad spectrum of subject areas, despite the fact that its holdings might appear to be relevant only to one. This emphasises that it is not obvious, sometimes even to collection providers themselves, what audience will appreciate the collection contents. Image collection provider case studies

3.1.3 Introduction

The scoping study commissioned TASI to identify a number of image collections from the collection matrix, and to provide short case studies of them in order to help us understand these collections, their needs, and how they might relate to any proposed model of a network of community-led collections.

3.1.4 Process and methodology for selecting case studies

We determined that case studies should:

- be 'good examples' of an image collection
- have particular features or functions that are relevant to CLIC
- be drawn from a variety of subject domains.

The matrix provided about 60 image collections that were particularly worthy of further research and investigation.

We then introduced more selection criteria to whittle down this number to 10–15 collections of which individual case studies would be undertaken. The selection criteria for stage 2 were representation:

- across the different academic disciplines
- from different funding bodies (e.g. HE and non-HE)
- of collections created by amateurs or enthusiasts, which are clearly of use within HE/FE.

3.1.5 List of case studies

For the completed case studies, please see Appendix 11.3.

Imperial War Museum – Concise Art Collection
http://vads.ahds.ac.uk/collections/IWM.html

Bioscience ImageBank
http://www.bioscience.heacademy.ac.uk/imagebank/

DoITPoMs: Micrograph Library
http://www.msm.cam.ac.uk/doitpoms/miclib/index.php

Art and Architecture
http://www.artandarchitecture.org.uk/

Gathering the Jewels: Welsh Cultural Heritage
http://www.gtj.org.uk/

Ingenious: Science & Technology
http://www.ingenious.org.uk/

Images of England
http://www.imagesofengland.org.uk/

Stone Pages Archaeology
http://www.stonepages.com/

CALVisual: Engineering Science
http://calvisual.lboro.ac.uk/

UK Moths
http://ukmoths.org.uk/index.php

Digital Egypt
http://www.digitalegypt.ucl.ac.uk//Welcome.html

British 20th century cartoon drawings
Collection database: http://opal.ukc.ac.uk/catalogue/ccc.pl
Collection homepage: http://library.kent.ac.uk/cartoons/

The CLIC study interviewed the large national image collections, SCRAN (http://www.scran.ac.uk/) and the Wellcome Trust Medical Photographic Library (http://medphoto.wellcome.ac.uk/). Both have considerable image assets numbering hundreds of thousands and have long experience of providing a targeted service to users. They were considered to be excellent examples of targeted image provision with particularly strengths in mature, easy-to-use interfaces to their collections and services. The CLIC study also discussed the history and evolution of the Biomedical image collection of 8000 images based at Bristol University with Medical HEA subject centre staff (http://www.brisbio.ac.uk/index.html).

3.2 Benefits of subject-based, community-led collections

The Rights and Rewards academic survey (Oppenheim, 2006) found that people generally turn to their nearest neighbours when looking for support materials and that the most active support is given to subject-based repositories rather than generic or institutional repositories. CLIC found that image providers showed a similar level of support for donating material to repositories.

Most of the more successful image collections, such as the BioScience Imagebank, and Medical image collections, such as the Wellcome Trust's Medical Photographic Library, are based on specific-subject disciplines; even SCRAN was originally focused on Scottish culture. Throughout its study, CLIC found that subject-based and community collections have had some exposure already, within relevant subject fields, and are better known than generic collections or collections based around subjects that differ from the respondent's own expertise.

Additionally, we found during the project that image users identified very strongly with their own subject area, and that all subject areas considered (rightly or wrongly) their own image needs to be unique.

With this in mind, we feel that it is important that any model supports communities to organise and help themselves and is, as far as possible, open-access. Key factors that support choosing a subject-centred model include:

- increased likelihood of community submission and sharing
- increased relevance of material and appropriateness of metadata
- ownership and trust
- common goals and values
- long-term sustainability
- existing archival activities
- existing funding structures and strands.

In order for a subject-based model such as we propose to succeed, it must contain material that is useful to the communities it is intended to serve. This can be

achieved by incorporating material created by a community of practice into a collection from the outset.

4 Investigation of user needs

CLIC has investigated the needs of image users and providers through:

- User Needs Conference (Oxford, 7–8 July 2005)
- Survey of attendees at the 2005 Gordon International Conference on Visualisation in Science & Education, (http://community.middlebury.edu/~grc/)
- Analysis of CLIC user needs from the Digital Picture Survey results
- Collaboration with projects in the JISC Repositories strand
- Survey of Image providers
- Interviews with image users and subject representatives

4.1 CLIC User Needs Conference

Over 40 people attended the User Needs conference in Oxford, in July 2005. They were either current or potential image collection owners, ranging from academics with large collections to librarians and technologists looking for solutions to collection needs. The primary incentive for people to attend was a workshop giving training on project management for building a departmental image collection. The conference was over-subscribed within two weeks of advertising. A series of talks about their own resources was given by collection owners; there were presentations on digitisation issues, social networks for image sharing (such as Flickr and the Internet Archive) and copyright and licence issues, including a talk on Creative Commons licences given by one of the team involved in localising the CC licences to UK law. All the attendees completed an online survey before the conference, reviewing their needs as users and potential collection owners.

We ran three focus group sessions:

Thursday:	User needs group session
Thursday:	Future needs open session
Friday: session	Attitudes to a national network of shared collections group

4.1.1 User needs group session

Attendees were split into two groups: those who had indicated in the pre-conference questionnaire that they were image collection owners or maintainers and those that were general users. Each group, comprising six or seven people, was asked to discuss a different topic. Each appointed a chair and a recorder/scribe, whilst a silent observer from the CLIC team also took notes.

The session looked at the following themes:

- Discovery of image collections
- IP and other rights
- Contextual information
- Sharing and community needs
- Sustainability
- Attitudes to models for a network of community collections.

During the session each group agreed on a number of comments. These are summarised below.

Theme: Discovery of image collections

‘We look for works in our own collections. If we don't have it we make it. We discussed how categories created by collection owners correlate with collection users e.g. museum catalogue is generally communicated to the public in an unmediated way. What's holding back digitisation – the thorny issue of rights and the sheer scale of the task.’

Theme: Intellectual Property and other rights

‘Who owns the image is a key question? Would they be happy for others to use them – probably yes to non-commercial purposes, but not for inappropriate use. There was a discussion on watermarking and using small versions to protect rights and the concerns around alteration of images.’

Theme: Contextual Information

“We need a minimum of contextual information if the collection is to be useful to us. Contextual information varies considerably from collection to collection and contributed information from users may indeed be more useful than original contextual information. There is information we would not wish to share due to issues of security, e.g. medical images.”

Theme: Sharing and community needs

‘Sometimes we're users of the images and suppliers of images – not really bothered about attribution, we all felt we should be sharing images if we'd created them ourselves. There is a conflict: feeling that we should be making stuff available for others, but institutions don't always want us to make material public. If image work is funded, each funder has different conditions that make sharing difficult, and a reminder that photographs could have multiple uses within different subject areas.’

Theme: Sustainability

‘We all have growing collections of material, it is important to realise they are not static. There are preservation issues, and we need a migration path to cope. There is a real need for standards to prevent obsolescence and for sharing out material.’

Theme: Attitude to models

‘The challenge is accessibility of images with diversity of sources, tracking down images. We need to decide what community we are talking about and where is the community? Widening participation is important in all of this. Support is important. If we can make things easier to do ... then this is important to the success of these models.’

‘Future needs’ open session

An open session was held to look at issues such as:

- intellectual property issues
- peer-to-peer file sharing
- the transition from analogue to digital technologies
- preservation worries
- the growth of social image sharing.

4.1.2 Attitudes to a national network of shared collections

These group sessions looked at the following themes on a positive or negative basis:

- dystopian view of a network of low quality shared material
- utopian view of a network of high quality shared material

Selected quotes from each discussion are presented below.

A dystopian view of sharing material in a community network

'Very similar idea to JORUM, many of the same facilities are suggested.'

'JORUM is fantastic idea, but institutions are 'snotty' about sharing resources. IPR and institutional resistance are the barriers.'

'JORUM provides cataloguers, which makes time less of a barrier than it normally is.'

'Cataloguers ensure that metadata is high quality, but can't ensure quality of the actual resources.'

'Institutional resistance is a very big barrier. Might need to link contribution to this type of repository to funding e.g. all CETL outputs need to go into JORUM. Educational and cultural value has to be judged by the content provider – don't want to end up with loads of committees deciding what goes in!'

A utopian view of sharing material in a community network

'You can't expect people to read whole licences every time they want to download an image. This is an interface issue: i.e. making the criteria for use identifiable on the screen [using icons/dividing up the screen] … Our suggestion is for a portal that gives access to images and tells you the criteria governing the use of each image. In an HE setting, the more open [the access], the better. [Short discussion about the importance of screen design.]'

A utopian view of sharing material in a community network from a negative standpoint

'Academic community: links go wider than the UK. Communities are international – would this suit a UK image library collection? We also need to take into account remote campuses abroad.'

'Quality material is expensive and academic institutions are competitors as well as communities – they may want to recoup some of the costs of digitisation rather than just simply share. Digital images are assets for the institution at the end of the day. Also, if establishments have put money into digitising collections they are going to be reluctant about being controlled by a higher body.'

A negative view of a sharing model

'It was suggested that the response to a single 'official' metadata standard would be to "b****r off", because ultimately it should be the user who establishes what metadata needs to be recorded. The value of an image is reliant on the metadata standard that it follows. Individuals disagree on what metadata is relevant, and therefore resent being forced to follow one strict metadata standard.'

'The problem is in metadata creation.'

'When creating new collections it is not that much of a problem; however, it is in the process of digitising old, established collections which is problematic. Furthermore, there is a difficulty in gauging who – and what type of person – will use the collection, which invariably influences what metadata is deemed necessary.'

'Funding issues create ad-hoc collections that only cover limited subject areas, [prompting discussion of the 'critical mass' issue].'

‘Subscription fees remain a problem, and create a disparity of usage.’

‘However, cost is inevitable with any collection, therefore it needs to generate money somehow.’

‘Perhaps a selective fee, where an individual pays on an image-by-image basis rather than a generic subscription charge should be considered.’

‘Issues of value for money (services must be user-targeted) when thinking about subscribing to services.’

‘Sometimes a condition of the initial funding prohibits making money, which thus makes long-term funding for maintenance very difficult.’

‘People do not generally complain about submitting images for a national-level database.’

‘Will some people be less inclined to spend time adding metadata to an image if the image will become copyright-cleared and not earn them any potential rewards? National-level negotiation gives excellent leverage in securing subscriptions and continual funding, while institutions and smaller collections often struggle. Would there be issues of prestige, publicity or something else?’

‘Image and metadata quality must be part of any national collection’s initial remit in order to ensure the long-term sustainability and use of the image collection.’

‘Would a ‘league table’ of image collections, where individuals may view which collection has the most images/users, serve any use? Or would such a table only serve to further stratify large and smaller collections?’

4.2 Survey of attendees at the 2005 Gordon International Conference on Visualisation in Science and Education

CLIC met attendees during a poster session at the Gordon International Conference on Visualisation in Science and Education in Oxford (http://community.middlebury.edu/~grc/). We discussed perspectives on sharing material and distributed a paper-based survey similar to the User Needs Conference attendee survey.

The survey covered general needs of image users attending the conference. The majority of the 28 respondents were used to sharing material and, as heavy users of image material, they were happy to accept material that was in a low-resolution format for teaching purposes. The session drew attention to the wide range of material that is freely available from American governmental agencies such as the NASA programme, including a wealth of geoscience image and multimedia material. This material can be freely reused with only attribution required. The results of this user survey are summarised in Appendix 11.7.

4.3 The Digital Picture Survey: an overview of arts education community needs

4.3.1 Introduction

This section is based upon the extensive research and consultation exercise carried out by the Digital Picture project, commissioned by the JISC Images Working Group and run by AHDS Visual Arts from April to September 2005. It draws upon the project’s final report, in turn based upon a national consultation of the community and all associated parties, using a questionnaire and focus groups. Over 500 individuals from over 150 institutions responded to the Digital Picture, including forty universities and over thirty associated museums. (See Appendix 11.8 for full report.)

4.3.2 Community involvement

The specific issue of community involvement was not raised by the Digital Picture questionnaire but several respondents used the 'general comments section' to stress how vital they felt it was. One noted that they needed 'Support for heterogeneous networks that allow teachers and researchers access to the hardware and software tools that are most suited to our needs, not those of the IT managers', whilst another put the issue particularly clearly: 'I'm pleased by your demand-driven, rather than supply-driven approach, since other initiatives don't seem to have learnt from the UK e-University mistakes.'

4.3.3 Intellectual property issues

Intellectual property was the subject of greatest concern to respondents. The community covered by the Digital Picture includes both creators and users of images, and responses to the survey embody the concerns of both groups: 68% acknowledge the need to protect the financial rights of image creators, whilst 75% believe the use of images should be free within education. There seems, therefore, to be little conflict between the two groups, with creators largely being willing to allow their work to be used for free for educational purposes – duly acknowledged.

The important issue, however, is what constitutes 'educational use': for example, one respondent noted that they would be happy for their images to be used in a lecture or school project without charge, but would expect remuneration for their use in a higher education course pack. This might be addressed by creating a generic licence (along the lines of the Creative Commons scheme) defining and allowing 'educational use'.

4.3.4 The Digital Picture's conclusions

There is a clear demand in the arts education community for a greater commitment to provision of digital images. However, such commitment must respond to the community's needs. There is a demand for community-led image repositories, and in order to meet the community's needs they must:

1. provide one-stop access to federated resources
2. incorporate locally-produced collections of images
3. allow institutions to house their own digital image collections
4. allow institutions to share image collections
5. allow users to deposit images
6. allow other public sector image collection owners to add their collections
7. have an acquisitions policy shaped by the community
8. create a critical mass of images relevant to the community's needs
9. contain a broad spread of material, avoiding the fine arts focus of most current repositories
10. incorporate existing public-domain images
11. aim wherever possible to acquire images of the highest possible quality, repurposing them as required
12. provide users with continuous online access
13. allow for serendipity
14. identify the level of quality and uses for which each image is suitable

15. ensure that images are provided with an agreed level of metadata, which must cover the images' identification, their provenance and their copyright status
16. define acceptable 'educational use'
17. try to acquire images with unrestricted rights for 'educational use'
18. state clearly and simply the copyright status and permitted uses for all images
19. secure sufficient funding and commitment from stakeholders.

This would best be implemented as a federated collection based upon harvesting and/or interoperability of discrete local collections, and would need the support of dedicated, long-term funding.

The Digital Picture has noted the JISC Image Working Group's proposals for community involvement and investment in facilitating and enabling future image provision, and has also proposed creation of a corpus of frequently-used images to help secure community support for these initiatives. In order to meet existing community concerns regarding quality, metadata, and intellectual property rights, it also proposes:

20. raising awareness of good digital image practice (e.g. quality, metadata standards, copyright)
21. encouraging image creators to make use of extended ownership mechanisms (e.g. Creative Commons)
22. ongoing negotiations with CLA and DACS to increase the potential provision of scanned images

5 The museum and commercial sectors

5.1 A survey of the provision of digital images by museums – summary of conclusions

Rupert Shepherd

12 January 2006

The full report in Appendix 11.9 is based on an examination of the museum-based online image collections listed, together with an examination of other leading museums and art galleries in the UK (particularly nationally-funded collections) and a few significant museums abroad. A full list, with URLs, is provided at the end of this report. Representatives of the following museums and museum projects were interviewed in greater detail from 1 to 14 December 2005 in order to establish the reasoning behind current provision and attitudes towards providing material for possible future community-led collections:

Bolton & Bury Treasures in Trust

COMPASS (The British Museum)

Ingenious (The National Museum of Science and Industry)

The National Galleries of Scotland

The National Gallery

The National Portrait Gallery

Tate

5.1.1 Conclusions

Image provision from the museum sector is top-down and generally aimed at an educated general public. There has been little consultation with academic communities regarding their image needs. Images are intended for personal and educational use, although restrictions are placed on their dissemination.

There is currently no established mechanism for museums and communities to communicate their requirements and concerns to each other. This can be rectified by the creation of a two-stage communication mechanism, based upon a community/museum liaison committee and a museum digitisation committee.

Museums are likely to be concerned about a number of factors, centred around loss of control over their data, the difficulty of clearing third-party copyrights in their objects, potential loss of revenue-generating opportunities, and the resources required to make collections available in a form that would satisfy academic communities.

These concerns should be met by the adoption of three basic policies:

1. Rather than create a single repository for digital images, aggregate data from individual image owners' collections
2. Adopt standard licences which address these concerns (most likely Creative Commons licences and one or two newly-created variations)
3. Provide funding to assist in the exposure of existing collections and the clearing of third-party rights in those collections

See the full report In Appendix 11.9.

5.2 Overview of commercial image collections

This section was compiled by Karla Youngs and Grant Young from TASI. Their report examines approximately 70 collections that cover a variety of business models. The report examines the following:

- usability, interface design
- metadata, search and retrieval
- control of access and delivery
- rights management
- charging and costing models
- community aspects

As might be expected both the quality of images and the systems used to deliver them vary considerably. The report highlights good practice in these systems' functionality, interface design, community building etc. It is interesting to see how sites that rely on their images for income deal with rights management and metadata provision. Some sites allow guests free access to their collections, others require registration and sometimes payment before anything more than thumbnails is delivered. Corbis is an example of a site that has both good metadata and a clear interface that allows users to retrieve images based on previous searches. Other sites offer 'mood' based searching, or provide ways of accessing frequently requested or random images.

There are many examples of innovative practice that could be adopted by educational collections in order to increase their appeal. The use of Creative Commons licences, image annotation, online communities and rating other people's submissions are all techniques that are described in the report.

The whole report is in Appendix 11.10.

6 JISC Repositories Strand collaborations

JISC is funding a large number of projects in their current Repositories Strand (http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories). Many of the projects are investigating technologies and ideas that are relevlant to the CLIC project. The following projects are of particular interest.

ASK: Accessing and Storing Knowledge – will develop a suite of open-source software that supports learners, researchers and teachers in securely accessing and sharing learning objects.

CHERRI-PIE: Common Healthcare Educational Recordings Reusability Infrastructure – Practice, Interoperability and Ethics – poses the question 'How can we protect patients, their right to absolute privacy while maximising the educational power of aggregated recordings and clinical stories?'

MIDESS: Management of Images in a Distributed Environment with Shared Services – will explore the management of digitised content (especially images) in an institutional and cross-institutional context through the development of a digital repository infrastructure.

RepoMMan: Repository Metadata & Management – will assist the development of repository infrastructure in several key areas by: assessing the feasibility of automated population of object metadata, conducting detailed user requirements analysis and review of associated digital rights management issues, adapting and providing a human interface to a generic workflow framework.

Rights and Rewards in Blended Institutional Repositories – will focus on the support issues, rights protection and rewards necessary to motivate teaching academics to use repositories and will blend the results with those required by research academics.

SPIRE: Secure Personal Institutional and Inter-Institutional Repository Environment – will focus on the setup and working through of the feasibility of peer-to-peer (P2P) technologies to aid the design of learning in the UK.

Trust DR Project – will explore in detail the emergent issues relating to the use of digital repositories within the UK HE/FE sectors, with particular reference to business processes and digital rights, from multiple stakeholder perspectives.

VERSIONS: Versions of Eprints – User Requirements Study and Investigation Of the Need for Standards – by reference to user needs and practices, will clarify researchers' requirements for deposit, storage and accessibility of different versions in the lifecycle of a digital resource.

6.1.1 ASK Project

The ASK project user needs analysis phase has short-listed a number of image repository applications that provide the functionality required by a wide range of practitioners.

The developers of these software systems have not generally considered implementing interoperability specifications and the advantages that compliance with frameworks such as the JISC Information Enviroment can bring.

The JISC Digital Repository and Supporting Digital Preservation and Asset Management in Institutions programmes are funding research into the key services that need building to provide an integrated set of repository services.

The CLIC project can be used to support the effort in defining these repository services where the emphasis is on using images in teaching, learning and research.

Images present a set of requirements for repository services that are different to the storage, discovery and sharing of academic artefacts, in particular:

- images cannot be described by text alone, users need to view images (or thumbnails)
- creating metadata that describes images is a very different process from the metadata that describes books, journals etc.
- images can be substantially larger in file size
- images can require specialist software to view and manipulate.

The purpose of the JISC Information Environment is to help developers create software that can communicate with other software performing similar tasks. The two programmes and effort by CETIS are helping to define exactly what is meant by each service and which interoperability specifications should be implemented.

As a summary of the services that are relevant to image repositories we provide this list:

Persistent unique identifiers: there has been sporadic but inconclusive effort to define a unique identifier schema over the last few years. The JISC repositories programme projects recently highlighted the issue again and tried to summarise and pull together the range of opinions and technical documentation (http://www.ukoln.ac.uk/repositories/digirep/index/Identifiers). Identifiers are needed so that users can point others towards a resource precisely and the discovery (or network location) of that resource can be separated from its metadata.

The Open Archive Initiative defines interfaces for harvesting or sharing metadata (by using web robots) between repositories. There has been discussion with the University of Southampton regarding how the network location of the image thumbnail could be transported with the metadata, thus allowing images to be shared more readily.

SRW, related to Z39.50 and SRU (see http://www.loc.gov/z3950/agency/zing/), allows federated search engines to find metadata pertaining to resources. There is no reason why this should not apply to images except the network location of the thumbnail would need to be part of the image metadata for similar reasons just stated above.

High-risk images: When searching for sensitive images (medical or other ethically-bound resources) there is a need for a user to be authorised to view and/or use an image. The work being carried out by OASIS on the XACML specification could allow federated search messages to contain personal data, which would mean only authorised people could gain access to metadata and thumbnails (in case of images).

6.1.2 CHERRI-PIE

The CHERRI (http://www.cherri.mvm.ed.ac.uk/) project identifies significant difficulties associated with the sharing and reuse of clinical recordings that need to be tackled at a national level. Clinical recordings are subject to patient consent, which must be gathered, and could, potentially, be withdrawn, as well as affected by privacy and ethical issues. This complicates any licensing scheme, and resulted in two key proposals that would need to be adopted by the UK healthcare profession as a whole.

CHERRI Consent and Licensing Model (C+LM)

The model requires a number of steps to be taken during the licensing process:

- consent evaluated, acquired and recorded along with recording
- consent encoded in licence and the licence + recording is passed for use
- licence stays with recording and outlines conditions under which it can be used (compliance dependent)
- licence can be matched with original patient and their consent, thereby establishing an audit trail.

The model requires a globally unique identifier linking the recording, its licence and the patient's consent agreement.

Clinical Commons

The CHERRI-PIE project proposed further policy work on the creation of a Clinical Commons licence scheme at the national level, which would allow the healthcare profession simply and easily to understand the uses to which the recording could be put, in a similar fashion to the Creative Commons licences. CHERRI's requirements for the Clinical Commons licences are:

- the four CC concepts of attribution, non-commercial, derivative works and share alike
- limits on the duration of the licence (e.g. until 31 December 2015)
- limits on the jurisdictions in which the licence applies (e.g. Royal Infirmary of Newtown only or UK only)
- limits on the scope of use (e.g. healthcare education only).

The last point implies that clinical recordings (CRANCS) should not be used outside the healthcare community. In CLIC's terms, clinical recordings are high-risk material and, as such, ought not to be incorporated into the three-tier model. Collections of high-risk material should remain within their user communities, and the CHERRI C+LM provides a workable model for how this could be done. CHERRI stresses that the implementation of their model will not be a trivial exercise and it is unlikely to be possible without significant effort and funding by stakeholders.

6.1.3 Rights and Rewards project

The Rights and Rewards study (Oppenheim, 2006) contains a lot of evidence to support the findings of the CLIC project. They surveyed 430 individuals from UK FE and HE institutions with the aim of gathering views on the use of an institutional repository for the deposit of teaching and learning materials. Two of the main areas of interest were:

1. What 'Rights' would individuals expect to exert over the teaching materials they deposit into a repository?
2. What 'Rewards' would motivate them to deposit their teaching materials?

Findings that are of interest to CLIC include:

- lack of awareness of national repositories
- strong alignment with and support for subject based repositories, with half of the respondents saying they would contribute to subject-based initiatives more than to any other kind (national or local)
- lack of awareness and confusion over IPR
- lack of guidelines and clarity in the labelling of resources, making people unsure of what they are allowed to do.

Image-based material is the most popular category of material that respondents would like to find in a repository; nearly 50% of respondents would contribute visual materials (a high score considering not all respondents are resource creators).

The biggest single reason for contributing was ‘to improve teaching’, the next ‘to support students’.

The biggest single reason for not contributing was lack of awareness of any repositories (43%).

The survey contains good information about how material in a repository should be licensed, the primary requirement being for attribution of the individual or institution. There is

reasonable mapping to the Creative Commons licences, though the key problem may be that an institution is unable to adopt a CC licence.

6.1.4 SPIRE: Peer-to-peer file sharing

The JISC-funded SPIRE project is researching the feasibility of peer-to-peer working across UK HE and FE institutions. SPIRE has chosen to investigate the LionShare system being developed by Penn State University specifically for academic peer-to-peer working. The LionShare site describes LionShare as follows:

> The LionShare P2P project is an effort to facilitate legitimate file-sharing among individuals and educational institutions around the world. By using Peer to Peer (P2P) technology and incorporating features such as authentication, directory servers, and owner controlled sharing of files, LionShare promises secure file-sharing capabilities for the easy exchange of image collections, video archives, large data collections, and other types of academic information. In addition to authenticated file-sharing capabilities, LionShare will also provide users with resources for organising, storing, and retrieving digital files.

Current use of peer-to-peer

We can only guess at the current state use of peer-to-peer software within UK HE and FE institutions, as peer-to-peer systems are officially banned from a typical university network. However, students do use peer-to-peer systems such as Skype and Kazaa, despite their institutions' policies.

In the public sphere, peer-to-peer is traditionally known as a method of illegally sharing media files. Peer-to-peer also has a reputation (deserved or not) among network administrators for using significant network resources.

Technical issues

The feasibility of LionShare in the UK is almost entirely governed by the method and implementation of the authentication system that it uses. At the time of writing, LionShare can only be used at institutions that have a Kerberos realm and a Shibboleth environment. Kerberos provides network security and Shibboleth the ability for users to take advantage of services in institutions other than their own. This choice of authentication method is forward looking and sensible but highly restrictive in the short term. Only larger institutions with specialist members of IT staff will be capable of setting up Kerberos and Shibboleth. LionShare alone is unlikely to be reason enough to do this. This effectively rules out FE institutions and any HE institutions which are not already planning to work with Shibboleth.

Policy issues

A typical university has a simple IT regulation relating to the use of ‘technology facilities’:

> (6) (a) No computer connected to the university network may be used to give any person who is not a member or employee of the University or its colleges access to any network services outside the department or college where that computer is situated……the unauthorised use of P2P software on machines connected the University Network is prohibited

This immediately negates the use of peer-to-peer networks that are connected by non-approved sources, particularly because it extends the use of the network to third parties. This is a major hurdle to overcome.

7 Possible technical solutions

7.1 Technical needs of an image collection

With reference to the collection provider survey detailed in Chapter 5, a typical community-led image collection has the following technical needs:

- Would like technical support and marketing assistance from a centralised service.
- Would like help in advanced searching techniques.
- Considers time, money and lack of technical knowledge to be the greatest barriers to future development of the collection.

It is important to remember that a typical image collection could be described as:

- built as a bespoke system running on a dedicated server in-house, providing open access via the web
- containing over 1000 images and expecting to grow to contain 10 000 images within two years
- lacking in advanced searching techniques, relying on thumbnail browsing
- providing information about items in its collection as Dublin Core metadata, if it provides metadata at all.

The key areas that should be addressed, in order to satisfy collection providers' needs, are marketing, searching, metadata and archiving. The mere fact that most systems are bespoke makes the requirement for technical support in these areas very difficult to satisfy. There was no consensus on the technologies used to implement image collections, with the exception that the vast majority were accessible via a web browser. This implies that solutions to collections' technical needs should be technology independent and web-based.

Even with the technical disparity between most image collections, there are some technologies that could be usefully employed to improve the current situation.

7.1.1 Marketing

Really Simple Syndication (RSS) is a simple alerting technology that is currently employed in a variety of areas, and across many different platforms. It has been adapted to serve diverse purposes, from calendaring to podcasting. There is currently an innovative method, promoted by the Flickr social image-sharing site, for describing an image in the content of an RSS entry, providing a thumbnail distinct from the image proper. RSS feeds could be used to promote popular image material and to broadcast it to interested parties and through aggregation, to institutional systems' lists of subject based material.

By adopting RSS for their sites' collection, providers could make real gains in marketing for a very small investment of time and effort. This is discussed further in Chapter 9.

7.1.2 Searching

Most image-collection sites do not have sophisticated search mechanisms. Collections are generally not held in database format but just as static html pages, and owners find that server-side searching functions are beyond their resources. Many rely on navigation through hierarchical menus, or on internet search engines such as Google to index their sites for them. None of the sites that we surveyed were part of systems such as Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

Collection providers could be encouraged to build upon the effort they have put into an RSS marketing service to help them provide metadata terms in the headers of their collections' web pages, which could be picked up by a centralised service, in a similar fashion to the

Creative Commons searches provided by Yahoo! and Google. The metadata provided would be similar to that required by the RSS marketing solution, so this would not place too great a burden on collection providers' time, and would provide an excellent stepping stone from RSS feeds to an OAI-PMH solution.

The OAI-PMH sets out a procedure for small 'static' repositories to be included in a Repository Gateway. Most of the collections that we surveyed are not 'static'; rather, they are growing at a reasonable speed, but it has been suggested that, in terms of OAI, a static repository is one that is not updated more frequently than once per day. The majority of community image collections would satisfy this criterion and could potentially be included in such a system. The apparent complexity of the OAI protocol puts it beyond the reach of most small collections but a centralised service could act as a Gateway to smaller collections. A long-term goal of any centralised service should be to combine search, metadata and marketing using SRW/OAI harvesting.

7.2 Use cases

It would be helpful to have technical use case scenarios for differently-sized community-led image collections with a variety of types of content. This would allow software designers to have a design specification or profile that actually addresses the needs of community image collections. Two very short examples follow.

Small *low-risk* CLIC owners appear to have the following needs

- Robust framework for exposing collections
- Archival facilities
- Search facilities
- Creation of formal meta-data structures

Small *high-risk* CLIC owners have the following additional needs

- Authentication controls
- Fine grained authorisation system for access and risk management
- Secure usage monitoring systems

The work done by CLIC, specifically the survey of collection providers, and the matrix of image collections, would serve well as a foundation for this work (see Appendices).

7.3 Review of software solutions

The project reviewed 18 pieces of software that are easily available in both the open-source and commercial sectors, and evaluated them in terms of the offered feature sets, mapping the needs of three hypothetical scenarios of small, medium and large collections. The suitability of each of the packages for connecting into an overarching network through metadata harvesting was also considered (See full review in Appendix 11.5).

Factors considered within the software packages included:

- Metadata: appropriate schemas (VRA Core 3.0, Dublin Core, UK LOM core, MODS), controlled vocabularies, taxonomies, folksonomies;
- IPR: Fair use, Creative Commons Licences, Copyright
- Resource discovery: portals, federated search (SRW), metadata harvesting (OAI), RSS
- Ease of installation, functionality, usability

7.3.1 Conclusion

Our survey found that community collections have four main requirements from image cataloguing software:

- content management
- online presentation with thumbnail and free text searching
- standards-based cataloguing
- federated searching.

Unfortunately there does not seem to be an immediately identifiable system that adequately covers the four areas, with federated searching being notably absent in most non-library orientated systems. Online content management systems do not provide much functionality for images, other than acting as a storage and retrieval service, lacking in customisable display and presentation aspects; however, they have advanced searching and retrieval services. Alternatively, popular packages for displaying images online are quick and easy to set up, offer many options for display and searching, but offer little in the way of cataloguing or content management and are ideally suited for small numbers of images with little descriptive metadata.

It is unlikely that any one piece of software will satisfy all user needs, and the fact that there is not even a 'market leader' is clear from the overwhelmingly widespread use of bespoke systems. If there is to be hope of centralised technical support for image collections, then promising pieces of software need to be championed and developed to make them more widely appealing. By targeting promising software that is already in widespread use (MDID is used by 40 American universities, Gallery is the most widely used photo display software, Index+ was the most-mentioned commercial package in our survey), and then adapting it to suit CLIC needs, it should be possible to achieve significant uptake of CLIC functionality in a short time.

7.3.2 Peer–to-peer software: applicability to CLIC

Given the technical implementation hurdles from which peer-to-peer networks suffer, it is unlikely that this technology will provide a suitable solution for CLIC-type network at any point in the immediate future, despite the fact that these solutions are targeted at letting academics collaborate in precisely the manner that would most benefit them. Large peer-to-peer networks will either be actively discouraged or require greater support from centralised institutional IT departments than is currently required by image collections. See Section 6.1.4.

7.4 Image metadata issues

The project investigated the importance of image metadata (EXIF, IPTC etc.) and appropriate schemas (VRA Core 3.0, Dublin Core, UK LOM core, MODS) and community annotation schemes such as controlled vocabularies, taxonomies and folksonomies. t is important to consider the distinction between individual image metadata and the metadata of the collection.

Metadata becomes important as image collections join up and become cross-searchable and federated. Metadata is also vital in discovery issues and aggregation of information for RSS channels. Unfortunately it seems that a significant number of collections do not hold much in the way of metadata about their collection or about the items in their collection.

Most collection providers seemed to be aware of the Dublin Core metadata scheme, and frequently state that their own schemas map to Dublin Core. Cataloguing was seen as an expensive and time-consuming activity for a collection provider to undertake and, given that time and money were among the largest barriers cited in our user survey, it is perhaps not surprising that this activity has been sidelined.

Very few of the collections that we surveyed present their metadata in a manner that would allow it to be used by harvesting or cross-searching mechanisms. Similarly, providers are hesitant to allow users to annotate their collections with comments.

7.4.1 Metadata needs of a small digitisation project

Metadata frameworks and generation

A project will create various forms of metadata, many of which can be generated automatically. Unfortunately, most of the information that is created automatically is not useful for discovery. This division is reflected in schemas such as the VRA3, which distinguishes between the image file and the object which it represents.

Typical Metadata types and appropriate schemas

Administrative

Image File properties – EXIF / IPTC / VRA3

Provenance

Creator, Subject – VRA3, Dublin Core

Linkage

Relationships between image files – VRA3

Subject – VRA3/Dublin Core/IPTC

Indexing

Image/Object relationship – VRA3

Rights Management

Image rights management – IPTC/VRA3 /Dublin Core

Dublin Core

A generic metadata standard aimed at producing catalogue records for a wide variety of objects, now a *de facto* interoperable standard. Dublin Core entries could be generated automatically from metadata files, or vice versa, since established mappings exist. This 15-field standard is often considered to be the key standard for linking diverse catalogue collections.

Exchangeable Image File Format (EXIF)

Exchangeable Image File Format (EXIF) is a technical standard for embedded image capture metadata and file-type metadata used by the electronics industry. In an image taken with a digital camera, the camera technical details, details on aperture and focus together with date and time information are automatically created by the camera and stored as EXIF information in the TIFF and JPEG image file headers. This should be captured automatically via either image processing software or the IMS. All good-quality image management systems should be able to take the EXIF information automatically. EXIF is system created information and it is not possible to place user annotation information here.

International Press Telecommunications Council (IPTC) header

Formally, the IPTC Information Interchange Model, now implemented as the IPTC Core Schema for the eXtensible Metadata Platform (XMP). IPTC holds generic image identification and rights metadata. Certain elements (e.g. Copyright Notice, Creator, Description, Keywords, Title) map directly to the Dublin Core metadata standard.

Visual Resources Association core categories, version 3.0

VRA3 allows for information to be stored about both the digital image and the object that it represents. It can also hold rights and file-type metadata. File-type metadata can be extracted from EXIF in file headers. Image identification and rights metadata can be extracted from EXIF and IPTC metadata in the image file headers. VRA3 maps directly to the Dublin Core metadata standard.

Collection-specific schemas

Significant numbers of collections that responded to our survey use their own metadata schemas to hold information specific to their own needs.

7.4.2 Post submission cataloguing of resources

A potential solution to the complexity and efficiency of cataloguing images by an author is the use of post-submission cataloguing techniques, where either a team of cataloguers or a user community add catalogue information or extra descriptive tags after the author has submitted. This has been found to increase the likelihood of submission of material to a system by reducing the form-filling hurdles on submission to the minimum necessary. Examples of post-submission proofing and cataloguing include Wikipedia, Jorum and Flickr.

User profiles and project profiles containing generic catalogue information may help but previous experiences at organisations such as SCRAN has shown that authors submitting material are only willing to enter minimal metadata; in fact, a rule of requesting a maximum of three fields only has been quoted by both JORUM and SCRAN.

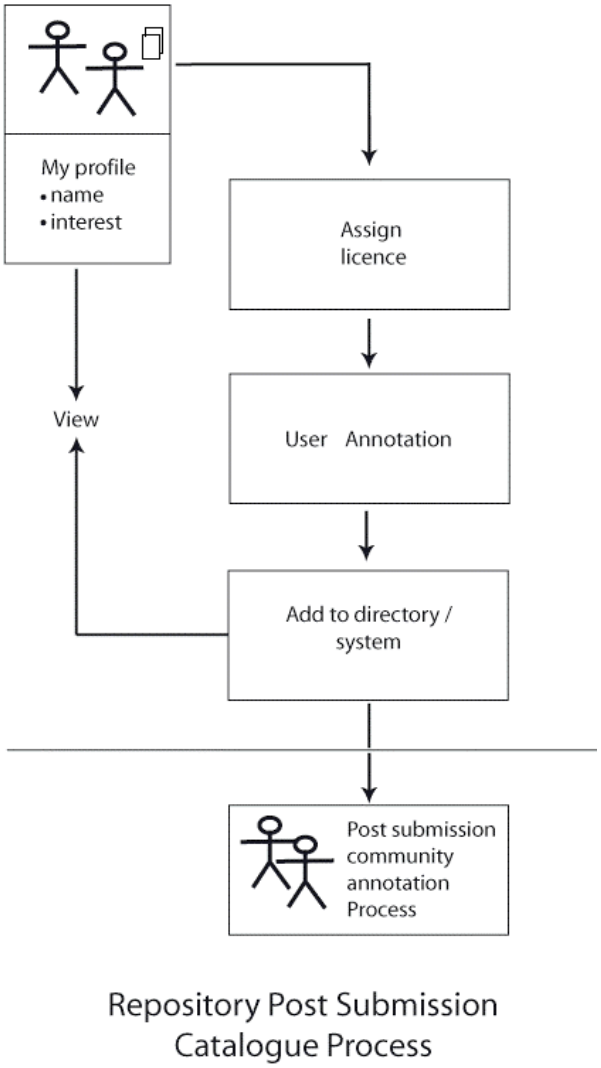

Figure 3 Community-based post submission cataloguing

Wikipedia

The Wikipedia online encyclopaedia project (http://wikipedia.org/) has shown the success of a post-submission cataloguing structure. It is now the world's largest encyclopaedia. Previous attempts by the same organisation to have a community submit entries to a free encyclopaedia, The Nupedia project (http://en.wikipedia.org/wiki/Nupedia/), failed to attract enough material, due to the complicated submission procedure. An extensive peer-review process designed to make its articles of a quality comparable to professional encyclopaedias characterised Nupedia. Nupedia wanted scholars to volunteer content for free. Before it ceased operating, Nupedia produced only 24 articles that completed its review process. Indeed the founder of Wikipedia and Nupedia, Jimmy Wales, admits he didn't realise the difficulty of the submission process to Nupedia and the barriers that were in place due to the complicated peer-review review process until he himself submitted an article. Similar experiences of authors being disenfranchised from submitting material to repositories by complicated submission procedures have been found in the eprint repository community. Although automatic alerts and verification mechanisms can be built into the system to help post submission cataloguing, Wikipedia still relies on a volunteer community to monitor the material submitted.

Most of the information in Wikipedia is submitted by ‘experts' in a particular field. Other users can contribute to pages and offer alternative views and information. Where a disagreement occurs, the topic is often 'branched' which allows multiple opinions to be expressed about the

same topic. Branching is controlled by the same community that monitors submission to the encyclopedia.

7.5 Discovery

7.5.1 Interfacing with national and institutional repositories

There are two traditional ways in which image collections could interact with national and institutional repositories, namely by metadata harvesting, or by federated searching. Both of these require considerable expertise to implement, and are likely to be beyond the scope of any but the larger collections surveyed.

Both methods would require standardisation of metadata used across image collections, which although not achievable instantly, would be possible given that most of the respondents to our survey recorded broadly similar information about their collections.

A new recent growth area of image discovery activity led by the social image sharers is that of 'Photostreams' or 'Photofeeds'. These are RSS 2.0 / Atom XML feeds linking to images. Photofeeds provide an easy standard way to reference a list of images with a title, a date and a description field. Lists of thumbnails of images with links to the main image are produced from a host image-sharing site such as Flickr per user or per publicly available selection. These feeds aid discovery as they allow a receiver, often a website using simple Javascript commands, to specify the type of image to display depending on the key fields metadata, size, date or other variable. Aggregation can also takes place based on grouping multiple feeds. Flickr also allows third-party tools and programming toolkits to access its infrastructure through a publicly documented API. There is considerable community-led activity in this area and much new innovative image tool work is happening. As the programming toolkits are becoming available, online java applets and FLASH objects are now being made by third-party users that provide niche interfaces to targeted searching, aggregation and filtering of huge amounts of online images stored in these public image banks. This will lead to a growth in available tools for the online manipulation and discovery of images.

These new innovations have prompted new ways to discover groups of relevant images for users and is appealing in situations where the user is looking for groups and types of images to choose from through visual thumbnail browsing rather than a specific textual search, The image feeds have much to recommend them, particularly if it is possible to aggregate on subject domain, keyword tags or specialist community nominated folksonomic tags. This RSS discovery technique would work well in interfacing community collections with national and institutional repositories and as a light-weight well understood and supported technology it is perhaps amenable to the smaller collections that lack sophisticated infrastructure but would like their material to be discovered more easily.

7.5.2 Discovery of image collections

Discovery of relevant image material was mentioned as a problem within focus groups at the July CLIC Workshop and in the Digital Picture survey. This is partly a discovery problem and partly due to the lack of social structures supporting educational image users. It is important that users of any proposed network or repository system have a variety of social forums, including mailing lists and bulletin boards for community discussion. These scaffolding structures will act to establish a sense of community, where devolved decision-making can take place.

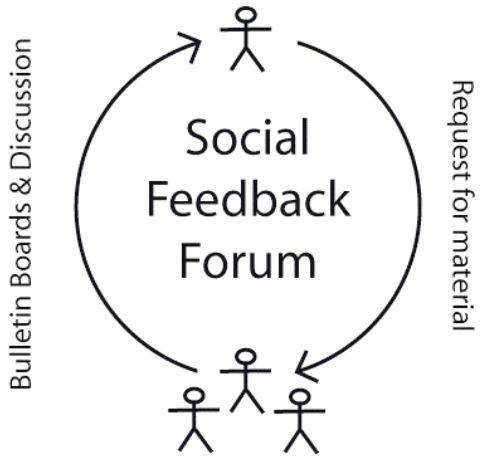

Figure 4 Social Feedback forum for repository work

Directory of image collections

There should be a nationally maintained list of UK publicly funded academic-related image collections held in an online directory. This directory should be maintained in an open way, with as few barriers as possible for people to register their material automatically. Checks should be made manually after submission by the list maintainers that the site is legitimate and relevant to the intended audience. Institutional webmasters should be alerted to the directory.

There is a need for consistent labelling of online academic image resources. It would be advantageous to standardise on metadata or textual strings that people can embed in their main index pages of their website, to flag that they have an image resource for others to use. A register of sites combined with consistent tagging of the sites would allow search robots to pick up the information automatically. A second stage of manually checking the results for validity and spoofing might then need to take place. However, at the moment there is, to our knowledge, no regularly maintained automatically generated directory of academic image resources. This is a low-cost solution to the discovery problem for users, with devolved image resources being discovered via a central register.

Image Portal: Central harvesting of online resource metadata.

The PIXUS demonstrator portal project (http://pixus.scran.ac.uk/) was an attempt in 2003 to link eight of the largest UK educational image collections together via the SCRAN service, using cross-searching of catalogue information. The project relied on the eight image collections adding or adapting cross-search architecture to their underlying delivery systems. Further work is continuing, under the JISC Portal funding strand, to investigate user needs for a full image and sound portal service. Anticipating this work in the future the CLIC project has tried to review how difficult it would be for smaller, ad-hoc community based collections, to be aligned with this work. What practical methods could these disparate collections adopt to expose their collections to a national portal, how expensive would the work be and would this be a viable sustainable solution? There are organisations, aligned with the Open Archive Initiative, that allow people to submit their archives and sites to a central registry of catalogue information. However, these sites are perhaps expecting a level of catalogue structure that most smaller community-led image collections don't have or, perhaps, don't have in the correct OAI XML format. There is a specification of how a smaller static collection of 5000 or fewer items could use a simplified version of the Open Archives Protocol for Metadata Harvesting but, again, this perhaps beyond the capabilities of owners of image collections that do not have specialist knowledge of this area (Specification for an OAI Static repository and OAI static repository gateway http://www.openarchives.org/OAI/2.0/guidelines-static-repository.htm). As expected, the survey of the providers found few of the smaller collections were built upon server applications that have in built OAI facilities.

Resource Discovery Network

The Resource Discovery Network (http://www.rdn.ac.uk/) has an important role to play as a reliable source of evaluated catalogue information on academic-related online resources. The RDN subject centres currently catalogue some information related to image collections, for example, they store 'image collection' as a category type. There may be variations in how these sites are labelled in the resource description: 'image bank', 'image collection', 'image archive' etc. and it may be useful to provide guidelines to clarify this for the cataloguers. RDN subject-centre networks should have 'image collections' as a search option, preferably on the front page of their websites. At the moment, only the arts-based network, Artifact does this, although other RDN centres have some image-related information in their catalogues that can be found via advanced searching. The Humbul and Artifact network has worked with the CLIC project to investigate advanced image-searching on their catalogue. A recommendation is that the RDN investigates any deficiencies in the cataloguing process and asks for image collections to have a prominent link on the main page of all subject centres and the hub RDN site.

HEA centre websites should also have mandatory links, perhaps using a standard pictorial icon, to their relevant RDN networks' image searching scripts, in their online resource directories. In a similar way all websites that are part of the JISC Information Environment should standardise on a pictorial icon to link to any proposed Community Image Repository.

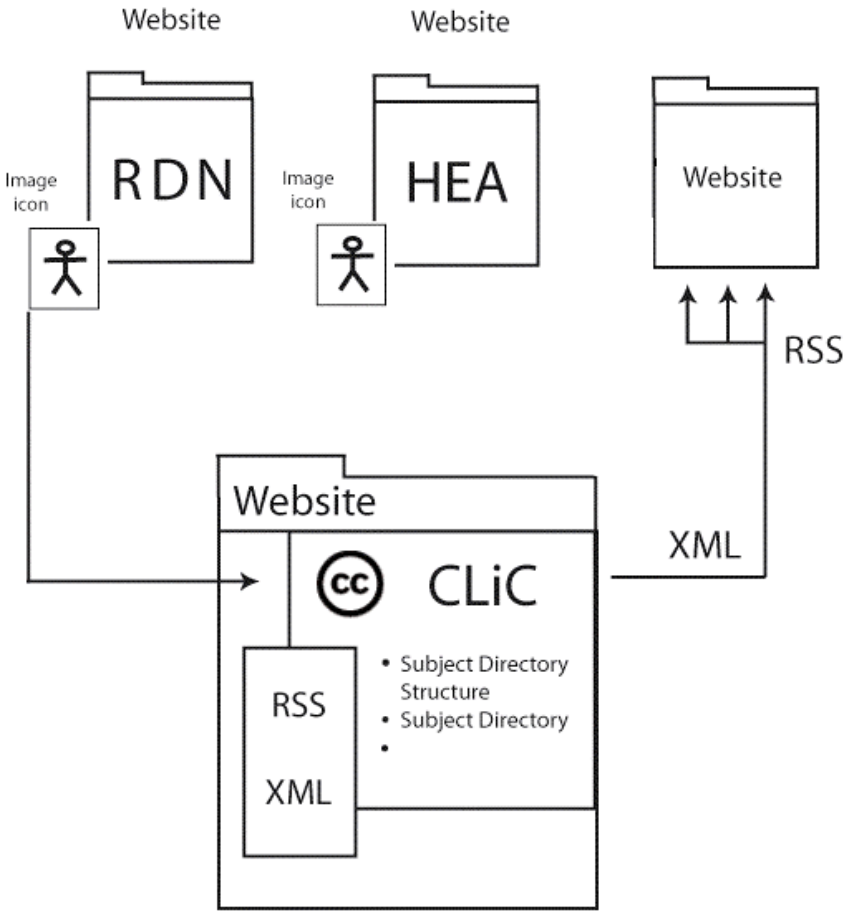

Figure 5 Clear labelling of sites in the three tier model promotes discovery of image material

Recommendations for the discovery of image material

- Any proposed network should offer opportunities for members of communities of practice to find material they need and discuss the material with fellow users.

- Any proposed network should increase opportunities for communities of practice to nominate their image material needs to higher funding bodies.
- Guidelines should be issued at national and institutional level, giving clarification of appropriate licences governing the reuse of image material.
- The Resource Discovery Network subject centres should expose their catalogue information on image collections prominently on their websites.
- The Higher Education Academy web sites should link both to the proposed central directory and the RDN image collections listings.

8 Rights and risks

Confusion over Intellectual Property Rights is seen by educational image users and consumers to be the main barrier to sharing and making better use of image material within UK HE teaching and research. There is widespread confusion and misunderstanding of the UK copyright laws, with the US concept of 'fair use' of materials wrongly ascribed to the UK copyright situation. Problems surrounding copyright were clearly the main discussion points in all focus groups at the CLIC workshop, the most frequently mentioned topic in conversations with image providers and users and this topic also dominated the Digital Picture survey. Image users need a simple way to discover material that is free for use in teaching and learning, and that carries with it a consistent and easy-to-understand set of usage licences.

Creative Commons offers authors a set of licences to assign to their own material, for instance allowing material to be reused for non-commercial purposes only and with proper attribution. Creative Archive is a similar licence model, created by a consortium of public sector groups such as the BBC and the BFI. Inspired by the work of Creative Commons, its licence stipulates further restrictions limiting reuse to within the UK, and allowing no endorsement and no derogatory use.

The Common Information Environment (CIE) group has submitted a report on the use of CC licences by publicly funded UK organisations. The CIE study attempts to clarify the risks and potential of making publicly funded digital resources available for reuse and investigates whether CC licences might reduce digital project costs and increase the visibility and reuse of such material. The report concludes that CC is suitable for some – but not all – of the cases studied. The Museums and Libraries Association also recently agreed to pursue the CC model.

The differentiation between institutional and individual ownership of image material has profound effects on how easy it is to reach agreement on sharing materials. An individual could make an informed, instant decision on releasing material into a repository for sharing; an institution obviously cannot. From the Jorum Digital Rights Management Watch 2005 report:

> *Lack of clarity and a common understanding regarding ownership of learning materials within institutions presents a challenge for Jorum. It may be expected that individuals will assume ownership of content that their employers would, if consulted, consider to be the property of the institution. If individuals were permitted to deposit content into Jorum, this may result in Jorum holding content with inaccurate (or disputed) rights information. It may be difficult to identify cases where the creator had incorrectly assumed ownership in ignorance. Jorum requires certainty of ownership. At present, Jorum accepts deposit only by institutions; individuals cannot deposit content directly. Thus, accuracy of rights information is the responsibility of the depositing institution; it must ensure that any staff member who is authorised to deposit has a clear understanding of ownership of rights in that content. However, this policy raises another issue. While content created in the course of employment is owned by the employer, not all learning and teaching materials used to deliver courses in UK HE/FE are owned by institutions as not all university and college teachers create their materials within the course of employment. For example, an employee who is paid an hourly rate and only for teaching hours creates the materials used for her classes outside of her employment contract and thus owns those materials.*
>
> *Nevertheless, the Jorum team believes that there are individuals within the community who, as rightful owners of the materials that they have created, wish to deposit those in Jorum. They could do this by assigning ownership of their materials to an institution to deposit on their behalf. However, a recent case in a*

Canadian University suggests that even when individuals are willing to share materials with colleagues, they wish to retain control over its integrity and use; assignment of ownership to an institution conflicts with that wish (Canadian Association of University Teachers, 2004). An alternative mechanism is available to allow individuals to deposit content in Jorum whilst retaining ownership; the owner may license her content to an institution within the JISC community and the institution may, in turn, deposit that content in Jorum (thus sub-licensing it to Jorum which, in turn, will sub-license it to Jorum users). This will allow Jorum to offer individually owned content whilst the institution bears the responsibility (and any liability arising) for it.

Whether institutions are willing to accept this liability remains to be seen.

Individuals working for institutions are often unaware of the IPR status of work (including image material) that they create in the course of their employment. Typically, this will belong to the institution that is employing the individual, but with funding for teaching and research frequently coming from different sources it can be very difficult for an individual to be sure who owns the IPR. Institutions and funding bodies should be more ready to donate material created under their auspices to the community, and this would best be achieved by adopting Creative Commons licences for any material funded by public money. There is evidence that a number of leading museums are seriously considering such a policy, but universities and other institutions will need to be lobbied to make them aware of the problem.

8.1 Risk

There are different risks attached to different kinds of image and associated contextual information. It would be useful to define this formally and have guidelines that can be adhered to by users and collection owners. There is confusion over acceptable use of image material, and privacy and ethical issues are also areas of concern. Issues to consider include:

- political
- ethical
- legal
- privacy
- misuse.

8.1.1 Recommendations related to risk

A rating system should be used to ascertain the risk factor of an image. Different models of CLIC network have different chances of success, depending on the risk factors associated with the material in the network. For instance, it may be considered that any supported open-access repository may not contain recognisable pictures of adults or children. Issues associated with high-risk holdings form the primary concern of the Clinical Recordings study, Cherri-Pie.

8.1.2 Medical images and clinical recordings

The CHERRI (http://www.cherri.mvm.ed.ac.uk/) project ran parallel to the CLIC study and identified significant difficulties associated with the sharing and reuse of clinical recordings, which need to be tackled at a national policy level. Clinical recordings are subject to patient consent, which must be gathered and could, potentially, be withdrawn, and also privacy and ethical issues. This complicates any licensing scheme for these resources and resulted in two key proposals of work that would need to be adopted by the UK healthcare profession as a whole before any clinical recordings material could be incorporated in any network.

A CHERRI Consent and Licensing Model (C+LM)

A Clinical Commons Licence

CHERRI Consent and Licensing Model (C+LM)

The model requires a number of steps to be taken during the licensing process:

- consent evaluated, acquired and recorded along with recording
- consent encoded in licence and the licence with the recording is passed for use
- licence stays with recording and outlines conditions under which it can be used (compliance-dependent)
- licence can be matched with original patient and their consent, thereby establishing an audit trail.

The model requires a globally unique identifier linking the recording, its licence and the patient's consent agreement.

Clinical Commons

The CHERRI-PIE project proposed further policy work on the creation of a Clinical Commons licence scheme at the national level, which would allow the healthcare profession simply and easily to understand the uses to which the recording could be put, in a similar fashion to the Creative Commons licences. CHERRI's requirements for the Clinical Commons licences are:

- the four CC concepts of attribution, non-commercial, derivative works and share alike
- limits on the duration of the licence (e.g. until 31 December 2015)
- limits on the jurisdictions in which the licence applies (e.g. Royal Infirmary of Newtown only or UK only)
- limits on the scope of use (e.g. healthcare education only).

The last point implies that clinical recordings (CRANCS) should not be used outside the healthcare community. In CLIC's terms, clinical recordings are high-risk material and, as such, ought not to be incorporated into the three-tier model. Collections of high-risk material should remain within their user communities, and the CHERRI C+LM provides a workable model for how this could be done. CHERRI stresses that the implementation of their model will not be a trivial exercise and it is unlikely to be possible without significant effort and funding by stakeholders.

Rights in repository technical work

In the repository domain there is a need to standardise on how licences are applied at the point of submission of an object. There is a requirement for recognition of different types of user needs, depending on the project. It is recommended that licences are not hard wired into the software but are treated as preferences per user profile or per image. Hence a contributor can choose and apply the correct licence at the point of submission. This is particularly important in cross-institutional activity.

For future repository projects, it must be flagged as a requirement that people exercise the option to assign Creative Commons or Creative Archive licences at the point of submission of digital material that they have created or are submitting. It should be made as easy as possible, by the submission interface, to determine an appropriate licence. An alerting system should be in place within the repository to notify administrators, after submission, to check whether the material can actually be assigned to these licences.

Discovery through consistent rights labelling

There is significant confusion over what an end user can do with material obtained from a website, or repository. Often the conditions for reuse within teaching and learning are not clearly stated.

The success of the Creative Commons icons for labelling online resources and the consistent use of associated licence hypertext has allowed search engines like Yahoo (http://search.yahoo.com/cc) and Google's Advanced search to return material licensed under agreed conditions (such as for non-commercial reuse).

There is scope for issuing best-practice guidelines for collection owners so that the benefits of this easy, automatic discovery of shareable content by search engines becomes well known.

Recommendations related to rights

- All JISC funded repositories should offer or support Creative Commons or Creative Archive Licence options on upload of material.
- Search for CC style material should be offered on all nationally funded repository collections.
- Guidelines for the use of CC licences should be issued for publicly funded image digitisation projects.
- Individuals in institutions should be given clear guidelines on the IPR status of their work.
- Guidelines should be issued to help collection owners understand the risks inherent in different types of material and the responsibilities associated with storing image material in publicly accessible collections.

Clinical recordings cannot be treated as normal images and require a separate licensing model to be constructed. The CHERRI-PIE project proposed further policy work on the creation of a Clinical Commons licence scheme at the national level, which would allow the healthcare profession to understand, simply and easily, the uses to which the recording could be put, in a similar fashion to the Creative Commons licences. CHERRI's provisional requirements for the Clinical Commons licences are:

- the four CC concepts of attribution, non-commercial, derivative works and share alike
- limits on the duration of the licence (e.g. until 31 December 2015)
- limits on the jurisdictions in which the licence applies (e.g. Royal Infirmary of Newtown only or UK only)
- limits on the scope of use (e.g. healthcare education only).

The last point implies that clinical recordings (CRANCS) should not be used outside the healthcare community. In CLIC's terms, clinical recordings are high-risk material and, as such, ought not to be incorporated into the three-tier model. Collections of high-risk material should remain within their user communities, and the CHERRI C+LM provides a workable model for how this could be done. CHERRI stresses that the implementation of their model will not be a trivial exercise and it is unlikely to be possible without significant effort and funding by stakeholders.

9 The CLIC model for a network of community-led image collections

9.1 Introduction

The CLIC project was set the task of researching and defining a suitable technical and organisational model to support the deposit and sharing of images by communities within Further and Higher Education. A model that supports the sharing of images amongst all possible users, from all subjects, each with its own particular interests and needs, is not feasible and we need to reduce the scope of the problem in order to develop a solution. The CLIC team believes that there are two sensible, achievable goals that should be pursued – increasing awareness of – and access to – image material provided by subject-based communities of practice.

Community collections are often provided and maintained with the intent of aiding learning and research, by image users who lack the technical knowledge and resources to provide more than simple open-access. Centralised support should be provided to promote sharing and exchange of material, bringing it to a wider audience. It is important to note that the promotion of sharing requires a primarily social, rather than a technical, solution.

9.2 The process

The first step in the CLIC project was to try to identify the extent of community-led image collections, through paper-based research and surveys. What collections were out there? How were they maintained? At what rate were they growing? Would they be amenable to joining a centralised network? This initial research was surprisingly difficult. Lack of previous research, lack of umbrella organisations and few directories for image collections, together with the transient nature of web-hosted collections, combined to make the initial task harder than perhaps it should have been and convinced us that there is a need for a framework to promote communication between interested parties.

TASI and the RDN were the only easily identifiable and accessible sources for directories of image collections. Both services also maintain contacts with image collection providers. The information they hold should be built upon with a series of registers, directories, social connections and mailing lists being set up, and maintained, as soon as possible. These directories should work to include material that is held outside of national and institutional collections, in what are often under-used and under-valued community collections. We anticipate that work to collate image collections will act as a goad to make these smaller collections known to the wider community.

9.2.1 Identifying existing collections

Using TASI's knowledge base, CLIC identified more than 400 open-access educational image collections holding high quality resources that are of value to the Further and Higher Education sectors. A good proportion of these could be said to be community-led collections with clear ownership by members of educational establishments or museums and libraries, and with the broad aim of increasing access to educational image resources.

9.3 Possible models

The CLIC project held a workshop with stakeholders, in which the following models for collection provision were discussed:

- national collections of images for each subject area
- a nationally maintained online subscription service of high quality licensed images
- a national online system or portal of catalogue information, linking to local collections

- a shared store of copyright-cleared images provided by members of the educational community
- a support network for autonomous local collections (e.g. providing technical support, directory services).

The intention was to identify the three most popular. There was little support for a subscription service, and image providers were not keen to donate their images directly to a central repository. Image collection providers felt that a national service might marginalise them, particularly if the system's interface removed the user from the original collection and its surrounding material. These all suggest that 'Disaggregation' (removing images from the holding collection) should not be attempted. Image providers want to retain control over images in their collections, and that any potential centralised system should seek only to refer viewers on to the images as they appear in their original collection.

We found most support for a central directory of image collections and for a system to hold catalogue information.

9.4 CLIC model

The CLIC model aims to satisfy the goals of increasing awareness of, and access to image material that is provided by subject-based communities of practice. The model takes a pragmatic, step-by-step view of linking up three key areas:

- institutional activity
- subject-oriented community-led collections
- national services, repositories and collections

9.4.1 Visual directory

The easiest, most achievable manner of raising awareness of image material would be to set up and maintain a 'visual directory' of image collections. This would be hosted centrally but maintained by subject specialists such as the Higher Education Academies. Collection owners would be offered a visual directory space offering useful tools, services and structures to registered collections. The kinds of service that could be offered include:

- directory information maintained by image provider
- simple provider profile information
- ability to host sample images from the collection in this space
- automatic news feed facility to alert users about new postings by providers
- technical support via mailing lists and bulletin boards
- aggregation of material and links from the providers' own associations
- marketing and access information
- standardised contact information.

The visual directory would be the starting point for a community-building exercise (see below), and will form the core of the CLIC three-tier model.

9.4.2 The CLIC three-tier model

The next stage would to create silos of sharable copyright-cleared material categorised by subject. Where possible, these silos should incorporate material that is held in community collections referenced in the visual directory, and also use material provided by national collections and institutional repositories. In the early stages each subject-based silo will need

to acquire material relevant to its subject where this cannot be obtained from existing collections.

The subject-based silos will need mechanisms for users to provide feedback on content, and make requests for material that is perceived to be missing. This should be linked to the visual directory, so that community collections can respond to requests for material, or make requests of national and institutional repositories. It is important to recognise that decisions about the type of material to be incorporated into a subject-based silo should be made by subject experts themselves using the community building tools provide by the visual directory.

9.4.3 Trust

During the CLIC study it has become clear that a major barrier to sharing of image material is a fundamental lack of trust between individuals, institutions and communities.

We feel that building a series of communities around a trusted third party such as a Higher Education Academy subject centre would nurture trust and cooperation between national, institutional and community collections. The trusted third party could reassure providers that material in other collections was relevant, high-quality, legal and fit for purpose.

- High-risk material should be held only in local repositories operated by communities of practice.
- Subject collections should hold only low-risk, rights-cleared material, and should be managed in a devolved fashion by subject experts drawn from subject based communities of practice.

Low-risk material held in open access subject collections should be shared with national collections, which should reciprocate by sharing their own holdings of low-risk, rights cleared material.

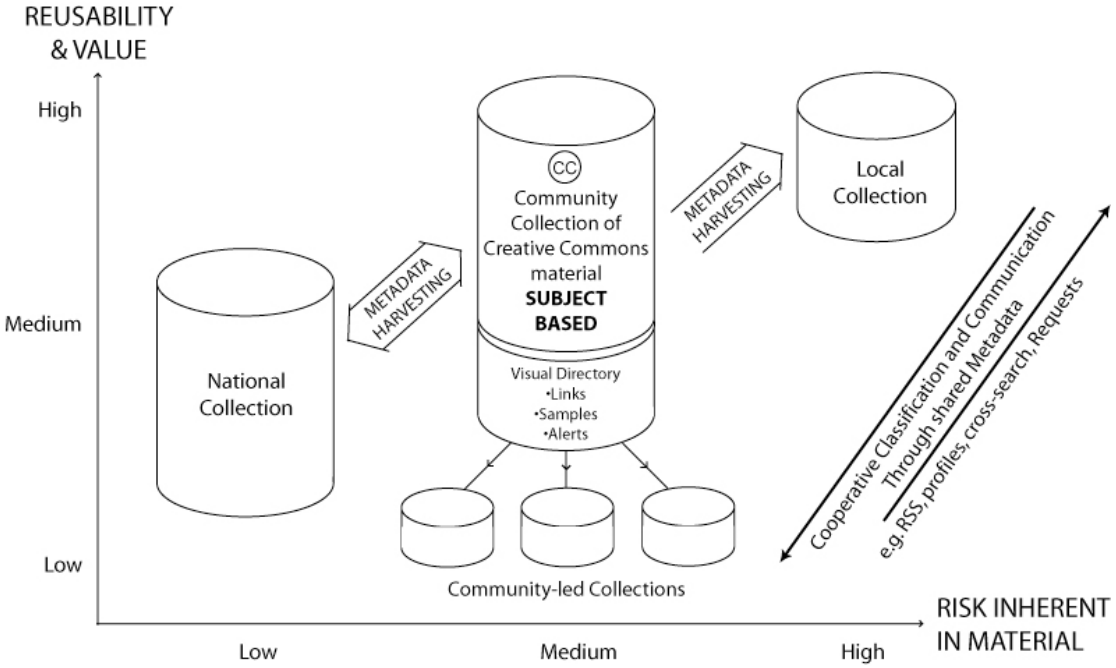

Figure 6 Proposed three-tier model for a network of Community Image Collections

9.5 Discovery and dissemination

The visual directory provides collection owners with a mechanism for marketing their collections and disseminating news via RSS photofeeds. These feeds should be picked up by subject-based sites such as the Higher Education Academies (HEAs), RDN, TASI,

National Grid for Learning (NGfL), British Educational Communications and Technology Agency (BECTA) and other centrally funded organisations with an interest in educational image material.

Each HEA centre and RDN hub should use an agreed icon, placed prominently on their website, to display links to subject-based image collections. Recent additions, popular search results and project news should be disseminated through Really Simple Syndication (RSS) photofeeds.

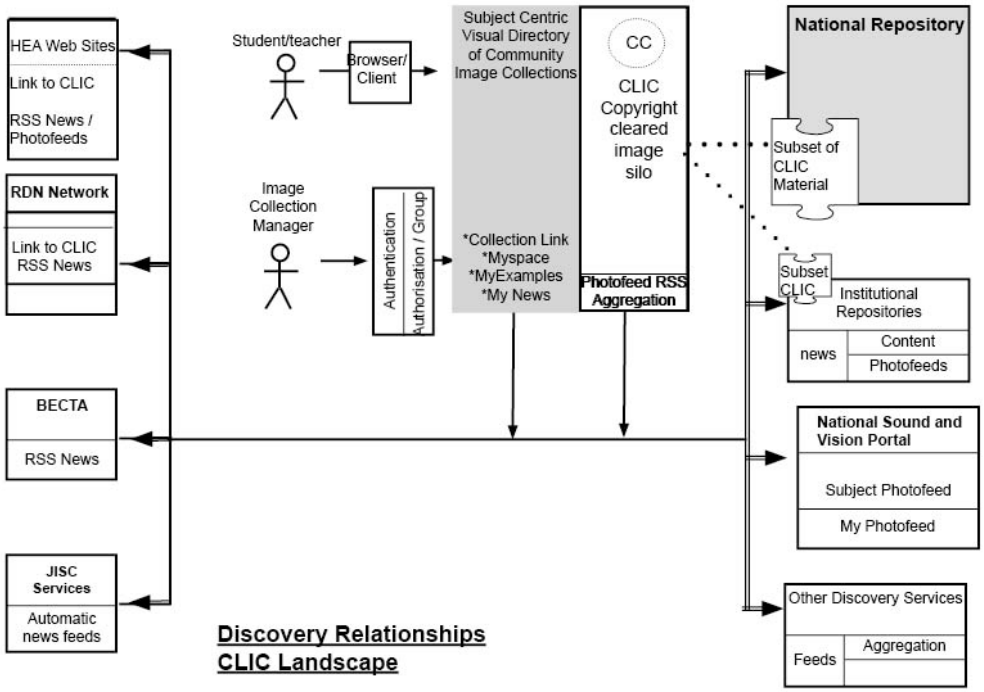

Figure 7 Discovery relationships between image providers and image seekers

9.6 Community building

The community aspect of the CLIC three-tier model is very important. Koper (R. Koper *et al*. 2004) proposed a model for the exchange of learning objects that suggests that it is important to construct certain use cases for building communities. We consider that images and their descriptive metadata are simple examples of learning objects, so it is worthwhile stating these use cases:

Key use cases to realise exchange from a technological perspective

- Find, get, edit and (dis)aggregate objects
- Upload, upload new and add metadata
- Feedback and logging of use
- Communication and collaboration space

Key use cases to realise exchange of material in a community

- Needs assessment
- Management and application of policies
    - IPR
    - Standardised terms and conditions of use

In order to build trust among community members, it is important from the outset to define roles for users within that community. There need to be *social moderators*, who are policy makers and upholders of the community's rules, and *facilitators*, who are active members with technical knowledge of the system and who provide support to other members and contributors. Both roles should be filled by members of the subject community who are aligned with their subject discipline and consider its needs from a subject-specific point of view. These users will already have knowledge of activity within their local sphere, department, research group, library or institution, broadening out to the national and international arena.

There are many excellent examples of community sites provided by the commercial sector and the features and functions offered by the best of them should be incorporated into the visual directory.

10 Implementation and cost-benefit analysis

10.1 Business and management model

The CLIC three-tier model will be implemented through a series of stages that promote social cohesion and grass roots support from the represented communities of practice. It seems sensible to aim for critical momentum as early as possible, building upon the foundation work of recent scoping study activity, and implementing the model in discrete, achievable stages.

10.2 Intellectual property rights

The success of the model relies on providing a subset of material central to specific subjects and, critically, material that does not need rights clearance. This material should be provided by the community but some may need to be funded specifically for the project. This material must be 'born digital', with no prior copyright, and licensed in perpetuity. Legacy legal and IPR issues have the potential to cripple development of the model in its early stages, as has happened to a number of the digitisation and resource creation schemes studied.

The success of the CC icons for labelling online resources and the consistent use of associated licence hypertext has allowed search engines such as Yahoo (http://search.yahoo.com/cc) and Google's Advanced search to return material licensed under agreed conditions (such as for non-commercial reuse). This approach will promote automatic discovery of sharable content by search engines.

Our investigations, and also the Digital Picture survey, revealed that there is significant confusion over what an end user can do with material obtained from a website or repository. The US concept of 'fair use' of materials is often wrongly ascribed to the UK copyright situation. Image users need a simple way to discover material that is free for use in teaching and learning, and that carries with it a consistent and easy-to-understand set of licences.

A policy decision should be made regarding the cost-effectiveness of developing a new licence to govern the sharing and reuse of image material. Considerable work has already been done by the Creative Commons community on localising their licences for the UK legal systems and this inspired The Creative Archive consortium to create UK-specific licences. Significant time and effort could be wasted on the drafting or adaptation of licences to address the perceived uniqueness of the educational field. If it is decided that existing licences are not suitable, OSS Watch and other relevant JISC legal working groups should be consulted early in the process to advise on alternatives. It is vital to achieve the goal of an initial set of material with absolutely clear guidelines on licensing, attribution, modification and sharing.

10.3 A visual directory as starting point

The community image providers have already shown considerable interest in – and support for – being associated with a national visual directory and there is similar support for a sharing of collections' catalogue information (See Appendix 11.4 Table 34). The CLIC project has already compiled a directory of 500 sites, containing an estimated 5–6 million images.

The CLIC image provider survey has also compiled a list of around 60 image collections that are willing to contribute to a visual directory with the survey providing preliminary subject classification work and a basic summary of the descriptive material and metadata that the participating collections keep.

The visual directory would be hosted centrally but maintained by subject specialists such as the Higher Education Academies. Collection owners who register would be offered a visual directory space providing useful tools, services and structures. The following kinds of service could be offered:

- A directory of image collection information maintained by image provider

- Simple provider/collection profile information
- The ability to host sample images from the collection in this space
- Automatic news/photofeed alerts created for every provider
- Technical support via mailing lists and bulletin boards
- Aggregation of material and links from the providers' own associations
- Marketing and access information
- Standardised contact information

The directory should hold a sample thumbnail gallery to give an overview of each collection and it is suggested that a template be provided for the collection to donate these small, low-resolution images, under a licence for reuse.

The visual system should act as an encyclopedia of educational image activity. There is a precedent for this idea – Wiki Commons (http://commons.wikimedia.org/wiki/) uses the freely available open-source *Wikimedia* software to generate an encyclopedia of images that can be reused throughout the Wikipedia project. A simple subject-based front-end and free-text search, together with recent addition searches, will aid discovery of image material. The maintenance of each collection page can easily be devolved to the partner image providers.

10.4 Support staff

The model requires very little in the way of central support staff, perhaps one or two people, with as much work as possible devolved to subject communities. This means time and effort should initially be spent fostering social networking and on face-to-face meetings of providers, themed along subject lines. If implemented in this fashion the costs should not cause long-term funding issues. The CLIC model will provide a valuable test-bed of image material and resources that could be incorporated into further national repository work, such as JORUM, institutional repositories and help aid the creation of e-Learning materials, and provide a use-case for portal work.

10.5 Roles and responsibilities

Key stakeholders:

- Image Related services TASI, AHDS Visual Arts, JISC Images Working Group
- Museum sector, Libraries and Heritage sector aka Common Information Environment
- HE Academy Subject Centres, e.g. HEA Art, Design & Media. Data Centres e.g. Archaeology
- RDN Network
- JISC Legal Advisory Team, OSS Watch
- National learning object repositories such as JORUM
- Interoperability and Metadata Standards bodies, e.g. CETIS

10.5.1 Roles and responsibilities of key stakeholders

Image-related services such as TASI and AHDS Visual Arts should give technical support and advice to help collection providers integrate into the model.

HEA subject centres and data centres such as Archaeology Data Service should maintain the visual directory and provide specialist knowledge from the subject perspective.

The museum and heritage sectors, libraries and the Common Information Environment should act as partners to the subject centres.

The JISC Images Working Group, JISC Legal Advisory Team, OSS Watch and Interoperability and Metadata Standards bodies such as CETIS should give guidance and direction to the subject centres.

The model requires very little in the way of central support staff, perhaps one or two people, with as much work as possible devolved to subject communities and stakeholders. This means time and effort should initially be spent fostering social networking and on face-to-face meetings of providers. If implemented in this fashion the model should not cause long-term funding issues; stakeholders should receive additional funding to support community image collections, and grants could be awarded to the community to encourage them to donate material.

In this way the CLIC model would provide a valuable test bed of image material and resources that could be incorporated into further national repository work such as Jorum and institutional repositories. It could help aid the creation of e-Learning materials and provide a use-case for portal work.

10.6 Implementation plan and cost benefit analysis

There are four key stages to the implementation of the CLIC three-tier model. These are:

1. social networking stage in which structures to support image sharing are formed
2. call for material in which subject-based silos are created, by issuing small grants to community image providers
3. catalogue-sharing stage: metadata is shared through photofeeds
4. national collections incorporate image-based material from the subject silos.

10.6.1 Stage 1 – Foundation or social networking stage

The foundation stage will use established communication methods and technologies to build communities among image providers.

- Social connections could be built up by means of an annual workshop or conference.
- Mailing lists to which interested parties can subscribe would be a low-cost way to gain access to individuals within institutions.
- Subject centres should be encouraged to provide information on the needs of their user communities.
- The Visual Directory should be expanded to become a “MySpace” Directory of image providers, and include sample material and a free RSS photofeed alerting service.

This work lays the foundations for social interaction, which is a necessity for any collaborative work. Activity in this area increases the potential for cross-sector discovery of material and would also provide a user-led forum for discussion of requirements on standardisation of licences, file formats, metadata interchange etc. It may also produce an amount of Creative Commons-style material if providers wish to showcase material from their collections. A visual directory that has a subject-oriented interface will also help gauge the extent of coverage across the various subject areas. The technical needs of the first stage are relatively lightweight and can easily be achieved with off-the-shelf, freely available open-source software:

- Wiki – Suitable for providing a visual directory with sample image material
- Blog Server – Suitable for providing news from the image provider sites
- Bulletin Board – Suitable for discussion of user needs and requirements

The visual directory would ideally be linked into the Information Environment Service Registry (IESR), which provides a method for computer systems, rather than humans, to

access the directory. This integration may have to be deferred until a later stage, when sufficient catalogue material is available.

This stage has clearly defined goals that are achievable within a short time frame and will build upon the work done for the CLIC study in identifying collections willing to participate in the visual directory.

It could be achieved with relatively small input from approximately two FTE staff.

Benefits – This work allows the foundation for social interaction, a necessity for any collaborative work. Any activity in this area increases the potential for cross-sector discovery of material. This activity will also provide a user-led forum for discussion of requirements on standardisation of licences, file formats, metadata interchange etc. It may also potentially prove a small amount of Creative Commons style material if providers wish to showcase sample material from their collections. A visual directory that has a subject-oriented interface will also help gauge the extent of coverage across the various subject domains.

10.6.2 Stage 2 – Subject domain call for material

Small funding initiatives would be used to create small teaching-based silos of copyright-cleared images. These should be managed by stakeholders, such as the HEAs who already award grants to members of their subject communities. Images collected under these schemes should be subject to certain requirements:

- All material must be born-digital or never previously published, and suitable for release under a CC-style licence.
- All images should have complete catalogue data that can be expressed according to Dublin Core.
- Images should be hosted by HEAs or other subject-based data centres, in a manner that would allow the images to be made available in Stage 3.

The technical needs of this second stage are lightweight but the processes of cataloguing and creation of descriptive metadata will be time-consuming and costly.

This stage requires a large amount of cross-sector activity:

- Short-term goal – Set up a call for material.
- Medium-term goal – Host material at HEAs or subject data centres.
- Long-term goal – Metadata can be harvested.

Benefits – Allows institutions and subject areas to target their internal digitisation processes with support from this core material

Cost could be comparable to a small software development team, perhaps 5 FTE.

10.6.3 Stage 3 – Catalogue sharing activity

This stage will bring together the dispersed image collections by making use of metadata that has been exposed in the preceding stage. Initially, this would be done via RSS photofeeds, but could lead on to an Open Archive Initiative Procedure for Metadata Harvesting (OAI-PMH) system that has been adapted to support images and multimedia.

Schemas already exist that would facilitate a pilot study to provide catalogue information expressed by means of Dublin Core and incorporate this information, together with a thumbnail image, into an RSS photofeed. These feeds could be aggregated by institutional consumers, subject-based portals and national services.

Opportunities would also arise for the host to post a submission cataloguing work by the subject communities of practice creating tags.

The OAI-PMH system would need to be adapted to provide the same functionality as is already possible with RSS and photofeeds but, in the longer term, it is likely that this modification would be of benefit to the OAI and image provider communities by enabling better integration.

This stage provides an excellent opportunity for the repository community to give guidance to image providers on best practice in post-submission cataloguing in various subject areas.

Even a little progress on cataloguing would allow better facilities for cross-search and discovery of image material. There is significant benefit in educating image providers in appropriate use of metadata; work to extend the OAI-PMH would benefit both communities.

This stage could be difficult to implement, due to a lack of underlying catalogue exposure mechanisms in community collections, and difficult to fund and coordinate on a large scale.

This stage is considered to be medium risk as it does not require a large initial outlay, but with a difficult to coordinate goal of sharing catalogue information across the model, it is likely to have a long ongoing time frame.

Benefits – Allows better cross-search and discovery. Any potential success here can be can be tied into National portal and learning object repository work. Simple progress on the catalogue work would allow subject aggregation of image material.

Drawbacks – It is difficult work to instigate, due to a lack of underlying catalogue exposure mechanisms in community collections, and perhaps difficult to fund and coordinate on a large scale.

10.6.4 Stage 4 – National collection activity

This final stage should be able to use images and catalogue material that has been produced in preceding stages. This material should be integrated with national initiatives such as portals and repositories that are currently being developed. Any proposed sound, image and multimedia portal, and learning object repositories such as Jorum, should aim to integrate with the CLIC model and incorporate material provided by communities into its own holdings.

This stage will also be difficult to implement. It involves coordinating activity across different sectors and performing technical interoperability work that may have little discernable result in the short term. If users' needs are to be met, there must be adequate consultation with the end user communities.

The success of this stage will rely on integration with:

- national portal activity
- learning object repositories such as Jorum
- national collections in the museum and library sectors.

10.6.5 Technical considerations

According to our survey, collection providers do not have the technical knowledge, time or money to implement up-coming technical solutions such as the OAI-PMH, Search/Retrieve via the Web (SRW) or Shibboleth. It is unlikely that collection providers will seek to implement these technologies unless they can demonstrate real value. It is much more likely that image providers would be able and willing to implement simple, well-understood proven technologies such as RSS photofeeds and Dublin Core metadata. The CLIC model seeks to build upon the adoption of these technologies to implement more sophisticated technical solutions currently under development.

The CLIC study identified archiving of content as a technical requirement that collection providers feel to be of value. Repositories and subject silos should consider offering an archival service as a way of encouraging image collections to contribute their material.

Similarly, the subject silos should implement solutions such as OAI-PMH and SRW for their holdings and actively encourage collection providers to implement the same technologies. This would best be achieved if software libraries implementing some of these technologies were developed in a variety of programming languages and maintained by advisory services such as TASI or AHDS Visual Arts.

10.6.6 Related risk assessment

The following table is an initial attempt at a risk assessment.

Risk	**Probability factor (*P*) (1–5)**	**Severity (*S*) (1–5)**	**Score (P x *S*)**	**Action to prevent/manage risk**
IPR delays on policy for non-clinical born digital material	4	5	20	Adopt a current licence model either CC or Creative Archives
Key stakeholders disagree on who should spearhead the initiatives	4	5	20	Concentrate on active subject areas in HE Academies
Collections do not join the visual directory	2	4	8	Encourage CLIC provider survey respondents to take part
Call for new rights cleared subject material does not illicit material	1	5	5	Adopt a funding strategy that instigates initial starting point
Subject HE academies do not wish to participate or do not have institutional infrastructure support	3	4	12	Fund key staff and resources centrally and encourage community involvement
Difficulty in compiling harvestable metadata from the image providers	4	4	16	Use RSS marketing strategy as a stepping stone to full OAI-PMH compliance
Lack of buy-in from the Museum and Libraries sector	4	3	12	Reinforce the message that an aggregated directory would mean that more users are led to the owners site
Danger of derogatory use of material	2	2	4	Clear policy on use

11 Appendices

11.1 References and Glossary

Arts and Humanities Data Service (http://ahds.ac.uk/)

Anderson, S. and Heery, R., 2005, Digital Repositories Review (http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf)

CHERRI-Pie, (http://www.cherri.mvm.ed.ac.uk/)

Common Information Environment (CIE) group, (http://www.common-info.org.uk) a report on the use of Creative Common licences across a range of UK publicly funded organisations (http://www.jisc.ac.uk/index.cfm?name=wg_cie_home)

Communities of Practice, Lave, J and Wenger, E. Situated Learning - Legitimate Peripheral Participation, Cambridge University Press, 1991

Creative Archive, a collaboration between the BBC, the British Film Institute, Channel 4 and the Open University. (http://creativearchive.bbc.co.uk/)

Creative Commons Licences (http://creativecommons.org/)

Digital Picture, the; a UK-wide initiative to explore digital image issues in the visual arts education community (http://www.thedigitalpicture.ac.uk/)

Dublin Core metadata initiative, provides a minimal set of metadata terms that can be used to categorise most types of material (http://dublincore.org/)

Enrich UK, gateway to lottery-funded online collections (http://www.enrichuk.net/)

EMBL, The EMBL Nucleotide Sequence Database constitutes Europe's primary nucleotide sequence resource (http://www.ebi.ac.uk/embl/)

Flickr, an online tool for storing, searching, and organising photographs (http://www.flickr.com/)

Gordon Conference, Visualisation in Science Education (http://community.middlebury.edu/~grc/)

Higher Education Academies, (http://www.heacademy.ac.uk/)

Information Environment Service Registry - a system to allow applications to discover and use materials which will help their users' learning, teaching and research. (http://iesr.ac.uk/)

Internet Archive (http://www.archive.org/)

JISC Digital Repository and Supporting Digital Preservation and Asset Management in Institutions programmes (http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories)

JISC Information Environment (http://www.jisc.ac.uk/index.cfm?name=ie_home)

Jorum – online repository service for teaching and support staff in UK Further and Higher Education Institutions (http://www.jorum.ac.uk/)

Koper, R et al. March 2004, ALT-J, Research in Learning Technology Vol. 12, No. 1, Building communities for the exchange of learning objects: theoretical foundations and requirements.

Museums, Libraries and Archives Council, (http://www.mla.gov.uk/)

Myspace, an online community site (http://www.myspace.com/)

Open Archives Initiative (http://www.openarchives.org/community/index.html)

Resource Discovery Network (http://www.rdn.ac.uk/)

Rights and Rewards study, Loughborough University Jan 2006 funded under the JISC Digital Repositories Programme (http://rightsandrewards.lboro.ac.uk/)

SCRAN, a long-established charity providing images and multimedia to education (http://www.scran.ac.uk/)

Search/Retrieve via Web, SRW, a system for returning search results from archives via web services (http://www.loc.gov/standards/sru/srw/)

Shibboleth, a standards-based, open source software package which provides Web Single SignOn (SSO) across or within organizational boundaries. (http://shibboleth.internet2.edu/)

Technical Advisory Service for Images, provides guidance and advice to the education sector, on the use of image material (http://www.tasi.ac.uk/)

Wikicommons, an online store of user-contributed, copyright-cleared media file for use in the wikipedia project (http://commons.wikimedia.org/)

Wikipedia, an online user-contributed encyclopedia (http://wikipedia.org/)

11.2 WP2 Matrix of UK Image collections

The matrix is available online at http://clic.oucs.ox.ac.uk/docs/WP2Matrix.xls

11.3 WP3 Image Collection Case Studies

Institute for Learning and Research Technology

CLIC Report from TASI

Report on Work Package 3

Document Notes

Authors: Karla Youngs and Grant Young
Date: 22 September 2005
Version: Final
Document: Name Report on Work Package 3 for the CLIC project

Contents

11.3.1 Introduction

This report covers WP3 undertaken by TASI and relates to the matrix (for WP2) that has been supplied to Oxford. It covers the following subject areas:

1. Process and Methodology for selecting case studies

2. Provides a number of mini case studies

For two, we have drawn together the information from the database and our own initial research into mini case studies. We recognise that further research, by Oxford, may need to occur in order to provide full case studies that are meaningful to CLIC.

11.3.2 Process and Methodology for Selecting Case Studies

In determining which image collections should be used to explore further the aims of CLIC, TASI adopted a two-stage process.

The first stage was to highlight collections that fit a number of criteria:

- Good examples of image collections
- Image collections that had particular features and/or functionality

Behind these selection criteria was the over-arching ambition to provide a list of potential case studies that cut across the subject domains.

From this initial investigation of the database, which held all data on the image collections, TASI was able to determine about 60 image collections that were particularly worthy of further research and investigation. This information was presented to the Oxford team in a highlighted (yellow) version of the matrix from WP 2.

Stage 2 of this process, then introduced futher selection criteria to whittle down 60-odd collections to about 10-15 collections that would be taken forward as case studies. The section criteria for stage 2 were:

- Representation across the different academic disciplines
- Representation from different funding bodies (e.g. HE and non-HE)
- Representation of collections created by amateur/enthusiasts which clearly have a use within HE

The aim was to achieve a good selection in terms of academic disciplines, funding bodies and 'official and amateur/enthusiast collections. This information was then presented to the Oxford team in a newly highlighted (yellow for broad selection, green for case studies) matrix.

11.3.3 Case studies

Case studies have been constructed with as much possible common information as possible, in order to allow further research to take place in WP 4 (non TASI work). Information that we have tried to document, includes:

From the database: Title; URL; Collection Description; Sector/Organisation; Subject; Funding Body; and Contact Name.

From the image collection's Web site additional information on: Copyright/Usage notices; Features that are unusual or standard in the delivery of the images; and any interesting points about the image collection.

Imperial War Museum – Concise Art Collection

http://vads.ahds.ac.uk/collections/IWM.html

This collection contains works of art from WWI and WWII, which aims to provide a social document of war with paintings from particularly important British artists. The digital collection has originated from within the Museum sector. The Imperial War Museum (IWM) is the holder of the original material, the IWM Concise Art Collection, and it is not indicated on either the IWM or AHDS-VA Web sites where the funding came from to digitise this collection. We suspect that the funding source was internal to the IWM, however this point does need clarification with the IWM. No one is named as a contact for this collection, users are pointed to a central email address art@iwm.org.uk For users who access this collection via its HE hosts, the Arts and Humanities Data Service – Visual Arts (AHDS-VA), a collection

description has been provided (see http://vads.ahds.ac.uk/collections/IWM.html) together with clear information on the images copyright/usage status within education. The IWM has cleared the copyright of these images for educational use.

When the collection was originally delivered via the AHDS-VA, the only access point for the collection within the museum was on a stand-alone terminal. However, subsequent to HE hosting, the IWM has now made the collection available from its own Web site at http://www.iwmcollections.org.uk/

The images made available to the education community through the AHDS-VA interface differ from other collections that are also hosted by the AHDS-VA: they are limited in size. Other collections make a range of sizes available for images e.g. if you click enough times you get a large high-resolution image. In contrast, the IWM Concise Arts Collection only makes the images available in thumbnail (for visual recognition) and 400 pixels along the longest edge.

This is a sizeable museum collection (2000 text and image records) delivered via the AHDS-VA for educational use. The reason for selection as a case study is that the AHDS-VA were originally the only hosts for the collection into the education community and subsequently the IWM is delivering the collection via their own Web site. As a follow up it would be interesting to know more about how this collection came to be delivered twice – with perspectives gained from both the AHDS-VA and the IWM. Another area for further exploration would be to investigate whether the AHDS-VA and the IWM developed a model for a museum/education partnership.

Subject areas that this collection cover are: HE - History; HE - History of Art, Architecture and Design; FE - Art and Design - Fine/Applied Art; FE – Humanities.

Clive Ruggles Image Collection

http://www.le.ac.uk/archaeology/rug/rug.html
http://www.le.ac.uk/archaeology/rug/image_collection/

This collection is composed of Professor Clive Ruggles' personal collection of photographs of archaeological and archaeoastronomical interest. He also manages the overall departmental collection at Leicester: http://www.le.ac.uk/archaeology/image_collection/ . No indication of who funded the collection is given, so we assume that Prof Ruggles has undertaken creating and building this collection in his own time with his own resources. Prof Ruggles is based within the HE sector and his collection has been built to support teaching and research needs. If anyone wants to contact Prof Ruggles, his email details are given on his Web site as rug@le.ac.uk

A copyright notice is given on the Web site and Professor Ruggles claims copyright ownership. After having found and displayed an image, each image comes with its own statement of permitted usage. The statement reads “You are free to copy and distribute this image to others in electronic form provided that (1) it is for educational use; (2) no charge is made; and (3) the copyright owner is duly acknowledged.”

For each collection a different search function has been offered. A text search function is available on the departmental collection, which has about 3000 images in it. A geographical browse function is available on Prof Ruggles own collection of 1100 images.

TASI has chosen this collection as a case study because it originated from an academic archaeologist who has own image collection, looks after his departmental collection and also hosts some student projects with images. Prof Ruggles is indeed a learning technologies and Image enthusiast.

Subject areas that this collection covers are: HE - Classics, Ancient History, Byzantine and Modern Greek Studies; HE – Archaeology.

Bioscience ImageBank

http://www.bioscience.heacademy.ac.uk/imagebank/

Bioscience ImageBank is a collection of teaching resources donated by academics, researchers, learned societies, industry and individuals. It was one of the first image collections in the UK to actively seek image donations from the Bioscience community, having 'cracked' the issue of metadata generation and the copyright issues surrounding image donation. As the image collection is hosted by the Higher Education Academy, it is firmly entrenched within the Higher Education community. The funding for this image collection has come from both the Higher Education Academy and the JISC. The contact for further details of this image collection is: imagebank@ltsnbio.leeds.ac.uk

Users can either search or browse for images. Before images are displayed a Copyright and Usage Page comes up. The usage notice is very explicit about what can and can't be done with the images and needs the user to click an 'acceptance' button (signifying they agree to work within the usage terms). If users cannot find an image that meets their needs, they can fill out a form requesting a particular image. ImageBank will then try to source an image to meet these needs. This indicates that the collection is being driven both by image donators and by image users in order to provide a relevant teaching resource.

This image collection is growing through community donation. ImageBank offer to undertake the digitisation activity for free in order to facilitate the donation of images. An area of the Web site has been made available for those who are donating their images electronically. A Web form has been created for image donors to fill in important information about the image and thus create the image's metadata. Information is also requested about who owns the copyright for the image being donated. Users of the image collection can also view images by contributor, so there is public recognition here for contributing to the collection.

Subject areas covered by the Bioscience ImageBank include: HE - Biological Sciences; FE - Sciences/Mathematics/Environment.

DoITPoMs: Micrograph Library

http://www.msm.cam.ac.uk/doitpoms/miclib/index.php

This specialist collection is composed of digitised versions of micrographs of a variety of materials. It was created specifically to enhance and support teaching and learning within the subject area of metallurgy. The collection has been created and is hosted by Cambridge University. No information is given on where or how funding was provided to create the image collection. The contact detail for the collection is a generic email address: doitpoms@msm.cam.ac.uk

The collection has a fairly simple interface and also has search and browse functionality. Users are presented with a 'Terms of Use' which explicitly states how the images can be used for educational purposes. Copyright for both the images and associated metadata, within the collection, is retained by the individual contributors. However there is also copyright for the collection database which is also made explicit.

This collection, like Bioscience ImageBank, also accepts donated images from donating academics. However unlike the ImageBank, there is no upload facility – donors must contact the image collection's manager to arrange contribution. Donated images must already be in a digital format for acceptance into the collection. Despite having strict image and metadata requirement the collection does have success in attracting image donations. As such the collection does not stay static but is growing.

Subject areas covered by this collection are: HE - Metallurgy and Materials.

Art and Architecture

http://www.artandarchitecture.org.uk/

This collection of images is derived from the Courtauld Institute of Art which falls within the categories of museum, HEI and image vendor all at the same time. The New Opportunities Fund supported the digitisation of the original materials. Deborah Swallow is given as the contact person for the collection: deborah.swallow@courtauld.ac.uk

The images have been organised into 'sets', which are all linked to from the home page. Users can also search for images by typing words into the search box on the home page. By doing so, this brings up a results page, from which the user can click on a thumbnail to bring up larger image plus basic metadata record. The opportunity to buy the image is presented underneath the image as either high quality print or digital file. However clicking on the image brings up a half-screen size image that can be right-clicked to download onto a user's hard drive.

Each image has its own copyright status – so copyright ownership is given as each image is displayed.

Once an image has been displayed, added value is given to the user through a zoom function. Users can draw a box on an area of interest in the image and the zoom function brings this up at full resolution. This is particularly useful for those subject areas which need to focus on fine detail of the image: e.g. brush stroke analysis.

Subject areas covered by this collection are: HE - History of Art, Architecture and Design; FE - Art and Design - Fine/Applied Art.

Gathering the Jewels

http://www.gtj.org.uk/

This is a very substantial digital collection that provides digitised versions of materials held within the libraries and archives of Wales. It has been created through collaboration of the National Library of Wales and the University of Cardiff with funding provided by the Big Lottery Fund. The contact for this collection is Leith Haarhoff at: loh@llgc.org.uk

The collection contains 20,000 digital images which have come from a very diverse range of original sources such as museum objects, aerial photographs, books and letters. The collection has been built to provide information about Welsh cultural heritage and is fully bilingual (Welsh/English).

Users are presented with a standard search box in which to type their keyword search terms. Users can also browse the collection within certain topic areas (which are presented on the home page) or browse through items of the collection which have an association with a particular Welsh town or village. Users are encouraged to provide feedback on many of the images – there is a feedback box under the image in which the user can supply additional information.

The images have been cleared for copyright when they are to be used for personal and educational use. However for commercial uses applications should be made to the museum, library or archive that holds the original object and the copyright for the digital object.

Subject areas covered by this collection are: HE - Archaeology; HE - History; HE - History of Art, Architecture and Design; Geography.

Ingenious

http://www.ingenious.org.uk/

This collection covers a huge range of subject areas, including medicine, transport, trade, science and technology, entertainment and media and the natural world. It has been made available through the collaboration of the Science Museum, the National Railway Museum, and the National Museum of Photography, Film and Television with funding from the New

Opportunities Fund. A generic email address is provided for those who wish to email: ingenious@nmsi.ac.uk

The collection has an innovative Web interface to a large set of images from various science and technology collections. The interface is very attractive with good navigation. The Web site features 30,000 images, 34 'articles' (which also have associated images so act like topic introductions) and several online debates.

The collection hosts a standard search box but also has advanced search functionality. This advance functionality can limit by category, can limit the search to a certain museum, or limit to specific images or the bibliography (which doesn't have full text).

The image collection is only available for personal, non-commercial use.

Subject areas covered by this collection are: HE - Hospital-based Clinical Subjects; HE - Art and Design; HE - Communication, Cultural and Media Studies.

Images of England

http://www.imagesofengland.org.uk/

Images of England is a large long-term project funded by the Heritage Lottery Fund and English Heritage. Its aim is to photograph every listed building in England as at 2001 (i.e. approximately 300,000 images). When we checked it had added about 175,000 images. If users wish to contact the English Heritage about the collection, they are presented with a Web email form: no direct email address is given.

The interesting thing about this project, and the reason TASI has suggested it, is that it is relies on volunteers to take the photographs. The volunteers are accorded copyright but grant English Heritage an exclusive licence to use the images. English Heritage provides its volunteer photographers with film, processing and travel expenses.

Images can be searched using a simple or advanced search, although registration is required for the latter. A recent addition (September 2005) is a 'Learning Zone', intended to enable students and teachers to use Images of England and other English Heritage image collections within the school classroom.

There is a lack of clarity around the copyright and permitted use of this collection. The terms and conditions (http://www.imagesofengland.org.uk/legal/tandc.aspx?pid=12) seem to make a distinction between the images (which are copyright the photographers and licensed to English Heritage) and the text (which is Crown copyright). The Crown copyright material may be used for teaching purposes, but no clear mention is made of whether the images can also be used for this purpose. The new Learning Zone enables such use but seems to be limited to schools.

Subject areas that this collection covers: HE - History of Art, Architecture and Design.

Stone Pages

http://www.stonepages.com/

Stone Pages is the work of two Italian-based enthusiasts, Paola Arosio & Diego Meozzi (info@stonepages.com), although a large proportion of the 529 prehistoric sites they document are from within the British Isles. Their collection is of interest to archaeologists, tourists and other enthusiasts and generates a lot of Web traffic and some fairly active discussion forums. Some of the "news" content is contributed by others, but the vast bulk of the content on the site has been created by Arosio and Meozzi. Their Web site was started in 1996 and has expanded to over 2000 pages.

Because of the organic way the collection has grown, navigation is sometimes confusing: 117 sites are accessed from 6 national collections (England, France, Ireland, Italy, Scotland, Wales), while a further 412 sites are covered in a series of tours

(http://www.stonepages.com/tours.html). Because the images are held within HTML pages, rather than a database, the search engine indexes and retrieves whole Web page links rather than images, but there are map and thumbnail-based browse options.

Some comments made on the site suggest that the authors have a lot of photographic content that is not yet available online because they have not had time to prepare it. They have introduced some commercial features (http://www.stonepages.com/shop.html) in an effort to fund further development. They have been successful in attracting sponsorship from some photographic companies and from SCRAN, for a project to document Scottish monuments. This project is available on the Stone Pages site (http://www.stonepages.com/ancient_scotland/home.htm), on CD-ROM, and presumably also via SCRAN's collection.

Online images tend to be small, but are sufficient for screen display. The copyright page provides permission for the pages to be printed or the images copied for educational use (http://www.stonepages.com/copyright.html).

TASI selected Stone Pages as an interesting case study because it is a personal collection that has developed into a resource of significant value to HE. Like the UK Moths collection (below), its authors maintain tight control over the resource and its presentation, however unlike the Moths collection, the owners of Stone Pages have not opened up their collection to other image contributors. From CLIC's point of view, the attempts at commercialisation and the sponsorship from SCRAN are also likely to be of interest.

Subject areas that this collection covers: HE – Archaeology.

CALVisual

http://calvisual.lboro.ac.uk/

This is an older project (1999-2000) funded under HEFCE's Teaching and Learning Programme, Phase 3, with the purpose of promoting the use of digital images within the learning and teaching of construction processes. The project was led by Loughborough University, but also drew on image collections from De Montfort University, University of Westminster, University of the West England, and Carillion Professional Services. The contact given on the Web site is the LTSN Engineering, which has now become the HEA Engineering Subject Centre (http://www.engsc.ac.uk/). The project manager, Dr Dino Bouchlaghem (n.m.bouchlaghem@lboro.ac.uk), is still at Loughborough and could be contacted.

Rather than mount the collection on a Web site, CAL Visual chose to package the image database as a CD-ROM, which was distributed to institutions, but can also be downloaded as a package from the Web site (600MB download). As the purpose of the project is to encourage lecturers to use digital images in their teaching, guides accompany the collection on effective creation and use of images and examples of how the CAL Visual images might be embedded within learning materials (see Prototypes).

We chose this collection because of its novel delivery method and because it contains some images and examples that are still useful in 2005. We note, however, that the quality of some of the images is poor and any repurposing of such a collection would need to be selective. This is likely to be an issue for many early image collections and may be something CLIC wishes to address.

Subject areas that this collection covers: HE - Civil Engineering.

UK Moths

http://ukmoths.org.uk/index.php

UK Moths is an example of a private collection built by an enthusiast that is very useful for academic purposes. Collections of this kind are common within the biological sciences and

are frequently linked to from academic pages (see, for example, http://www-biol.paisley.ac.uk/bioref/Animalia.html).

UK Moths currently contains 3427 photographs representing 1615 different species of moth (out of 2400 species recorded in the UK). The collection was started by Ian Kimber (ian@ukmoths.force9.co.uk) in 1998 using images taken from videos and, later, a 35mm camera. The site started off as hand-coded HTML on Geocities, but has migrated to a database (Access, then MySQL and PHP) on its own domain. UK Moths now also invites contributions from others and according to the 'About this site' page these now outnumber Kimber's own images. A page of 'Guidelines for Contributors' (http://ukmoths.org.uk/contguidelines.php) makes it clear what is expected from those wishing to add to the site. Kimber negotiates with potential contributors over the best way for them to deliver images and retains full editorial control, researching and writing the accompanying descriptive text himself.

Images displayed on the site are of screen-size resolution and are free to use for educational and non-commercial purposes as long as the author is notified. Higher resolution versions can be requested from Kimber or from the other contributors. The images are accompanied by scientific names and references, and a descriptive text, and they can be searched or browsed by species or thumbnail.

TASI selected this collection as a case study because it is a good example of a personal collection that has evolved into an important resource and developed a community of users and contributors around it.

Subject areas that this collection covers: HE - Biological Sciences.

Digital Egypt

http://www.digitalegypt.ucl.ac.uk//Welcome.html

Digital Egypt was a JISC-funded project funded under the Museum Content cluster of its 5/99 Learning and Teaching programme. A summary is available on the JISC Web site here: http://www.jisc.ac.uk/index.cfm?name=project_digital_egypt .

Led by Stephen Quirke (s.quirke@ucl.ac.uk), an archaeologist from the UCL Institute of Archaeology and curator of the Petrie Museum, the Digital Egypt project repurposed existing digital content from the museum (http://www.petrie.ucl.ac.uk/) for use as an HE learning resource. Images from the Petrie collection were combined with 3D virtual reconstructions produced by the university's Centre for Advanced Spatial Analysis (CASA) and with supporting materials (e.g. texts, maps, diagrams, tables) produced by Quirke and his colleagues.

Instead of a database of resources or a structured browse, Digital Egypt presents a variety of routes through the collection of images and information, typically placing them within a narrative context.

TASI believes this project would be interesting to investigate because of the way it re-use and contextualises existing heritage images and supplements these with new images (maps, 3D-VR) within a narrative framework.

Subject areas that this collection covers: HE - Geography; HE - Anthropology; HE - Classics, Ancient History, Byzantine and Modern Greek Studies; HE - Archaeology; HE - History; HE - History of Art, Architecture and Design.

British 20th century cartoon drawings

Collection database: http://opal.ukc.ac.uk/catalogue/ccc.pl Collection homepage: http://library.kent.ac.uk/cartoons/

This is a substantial online collection of cartoon images held by the Centre for the Study of Cartoons and Caricature at the University of Kent.

The online database contains details and images of approximately 90,000 20th century British cartoons (we note that another statement on the Web site says 35,000). Basic and advanced searches are offered, with look-ups for some of the metadata fields (artists names are alphabetised by forename; the keywords offer thesaurus functionality). Screen-sized images are delivered at 600 pixels long, with full images often very large in size (we saw one 5000 pixels long). Images are in the PNG format and are bitonal (black/white, no grey or colour), which enables efficient delivery but can compromise the quality of some of the images. As most of the works are still in copyright, users are instructed to seek permission before downloading or printing and a contact is given for those who want to advice on this or better quality reproductions. The email contact for the database is J Newton j.m.Newton@ukc.ac.uk .

We note that this existing collection has become part of a project held by the Research Support Libraries Programme (RSLP) to produce a 'cartoonhub', the details of which are available here: http://library.kent.ac.uk/cartoons/collections/cartoonhub.php . A collaboration between the Carton Centre and the LSE, University of Manchester and National Library of Wales, the cartoonhub project intended to digitise collections in each institution and enable them to be cross-searched from a central search on the cartoonhub Web site. This work was scheduled to end in 2002, but as yet there does not seem to be a cross-search facility.

TASI chose this collection as a potential case study because of its interesting format and collaborative/cross-search model. It would be interesting to learn about the status of the cartoonhub project and what difficulties they have encountered in developing a distributed collection.

Subject areas that this collection covers are: Subject: HE - Politics and International Studies; HE - History of Art, Architecture and Design.

11.4 WP 4 Survey of Image Collection Providers

11.4.1 Survey Results

Introduction

In liaison TASI, and after consulting with stakeholders and case study projects we constructed a survey of 50 themed questions designed to be delivered to the 500 community-based image projects identified in the matrix of collections (See Appendix 11.2 WP2). The survey investigated the attitudes of image collection providers to the obstacles and hurdles that they face in collection building, and any needs that could be met by national initiatives. The survey also asked questions about the access and rights that govern the use of their material, and the attitudes of collection owners to sharing their material with educational users. The survey gathered the following data:

- The intended audience of the repository, and the actual audience.
- The rate of growth of the repository, its perceived lifetime, and the size of its audience.
- Community-nurturing features such as commenting, forums, user additions.
- Access restrictions and rights management.
- Content monitoring and content quality.
- Obstacles faced during setup, and the technical solutions adopted.
- Needs that could be met by national models.
- Mapping of attitudes to sharing material with Creative Commons licences.

The survey was piloted by the digital image community at the University of Oxford.

The full survey took place from October 2005. See Appendix 11.4 for the full survey

Research questions the survey is intended to answer

- What missing technical infrastructure needs could be solved by a national CLIC service?
- What metadata exists locally that could be exposed to a national service (portal or directory)?
- What are the collection owner's attitudes to sharing and allowing open access for educational use?
- Is the collection part of a community? Is there a demand for sharing material?
- What are the key barriers to growth, and what are the technical support needs that could be addressed centrally?

The survey was devised and piloted with assistance from local (Oxford-based) images collection owners. Two pilots were undertaken, one using a paper version the other an online version. Comments received during the course of the pilots are discussed in the Pilot Surveys section below.

The questionnaire was distributed to the collection owners identified by TASI (see Appendix 11.2 WP 2). Many of the collections (280) provided only online contact details, so these were asked to fill in the online version of the survey. The 220 collections that provided postal addresses were sent the paper version of the questionnaire and stamped, addressed envelope for their response.

A total of 87 valid responses were recorded, although some respondents answered missed some questions there were no duplicate submissions.

Pilot Surveys

The pilots ran consecutively in October and November 2005. They brought to light a general lack of awareness of some of the technical terms used in the questionnaire and ambiguities surrounding the wording of questions around copyright and ownership of material. The questions provided were revised and more space was given for additional information to be given by respondents.

The pilot studies highlighted the need for background information on the CLIC project and its research aims, a formal data protection statement, and avoidance of jargon and technical terminology. The questionnaire was reduced in length and themed in to coherent sections, with the more difficult technical section moved to the end. It was hoped that if the respondent was not involved in the technical side of the collection we would still gain some useful information from the preceding sections.

Collection details

The first section of the questionnaire asked for information on the title of the image collection, the contact person and related institution or group, together with the collection website address. The image provider also completed a confidentiality statement.

Q6 Please describe your involvement with this image collection (tick any that apply)

	Response Percent	**Response Total**
I am the main contributor to the collection	14.5%	10
The collection is a major responsibility of my employment	36.2%	25
The collection is one of many responsibilities of my employment	**49.3%**	**34**
The collection is a sideline activity or hobby	10.1%	7
Other (please specify) *	13%	9
Total Respondents		**69**

Discussion: The majority of the respondents (49%) did not consider the image collection to be a major responsibility of their job but one of many responsibilities although 35 respondents were either the main contributor of material to the collection or considered the collection to be a main responsibility of the collection.

Q8 Please indicate the subject(s) to which your collection is relevant (you may tick more than one box).

	Response Percent	**Response Total**
Medicine and Dentistry	17.1%	12
Nursing, Anatomy, Physiology and Pathology	17.1%	12
	21.4%	15

Biological Sciences, Biology, Botany, Zoology and Psychology		
Veterinary Sciences, Agriculture, Forestry	10%	7
Physical Sciences, Chemistry, Materials Science, Physics and Geology	10%	7
Mathematical, Operational Research, Statistics and Computer Sciences	4.3%	3
Engineering	10%	7
Technologies, Metallurgy, Ceramics and Materials Technology	11.4%	8
Architecture, Building and Planning	37.1%	26
Social studies, Economics, Politics, Anthropology, Human and Social Geography	51.4%	36
Law	7.1%	5
Business and Administrative studies, Finance, Accounting, Marketing & Accounting	5.7%	4
Mass Communications and Documentation, Media Studies, Publishing and Journalism	11.4%	8
Linguistics, Classics, English studies and related subjects	18.6%	13
European Languages, Literature and related subjects	15.7%	11
Eastern, Asiatic, African, American and Australasian Languages and Literature	14.3%	10
Historical, Archaeology and Philosophical studies	**62.9%**	**44**
Creative Arts, Music, Cinema, Photography and Design	55.7%	39
Education	37.1%	26
Total Respondents		**70**

Discussion: The image collections covered the wide range of disciplines covered by the JACS subject coding. As expected there was considerable coverage of historical and political subject areas due to the large number of museum and heritage image collections that contained digitised historical material. The subject areas Mathematical and Business and Finance had the least coverage. It is to be noted that the image collection providers often selected multiple subject areas that they felt their image collection was of value to. It is important to note that a particularly large collection that covered a whole broad spectrum related to a subject like Medicine could be of use to a number of subjects such as history, design and technology and this relevance might be not obvious at first.

Q9 Please indicate the age range(s) to which your collection is relevant (you may tick more than one box).

	Response Percent	**Response Total**
Primary Education (4-11)	40%	28
Secondary Education (11-16)	67.1%	47
Further Education (16-21)	80%	56
Higher Education (University)	88.6%	62

Life-long learners/General public (Adults)	**91.4%**	**64**
Other (please specify) *	11.4%	8
	Total Respondents	**70**

Discussion: Most of the image collections (80% -88%) considered themselves to be appropriate to post-16 education, Life-long learners or the General Public. As 83% of the collections were open access web based systems (see Table 11) this suggests that most collections surveyed are arranged for a general educated audience.

Q10 What are the sources of the digital images in your collection (you may tick more than one box)?		
	Response Percent	**Response Total**
Existing collection of slides/photos	64.3%	45
Existing collection of digital images	41.4%	29
Photos/slides created specifically for the collection	35.7%	25
Digital photos/scans created specifically for the collection	**74.3%**	**52**
Other (please specify) *	22.9%	16
	Total Respondents	**70**

Discussion: A wide variety of material was represented. The image collections were often collections of digitised material (74.3%) with a reasonable amount (41.4%) using existing digital material. An interesting question would be to see how this changes in the next few years with the rise and spread of digital cameras and the expansion of consumer focused desktop image management systems. Other sources of material included early printed books, manuscripts, newspapers, printed directories and original artefacts.

Q11 If your collection contains any other media files, please indicate what these are:		
	Response Percent	**Response Total**
Audio (e.g. MP3, RealAudio)	60.9%	14
Video (e.g. MPEG, WMV, QuickTime)	**69.6%**	**16**
3D modelling (e.g. VRML, QuickTime VR)	26.1%	6
Other	17.4%	4
	Total Respondents	**23**

Discussion: Less than a third of the collections completed this question and in that group video was the most popular resource to store alongside the main image material, followed by audio then 3D modelling formats, other mentions included Macromedia Flash animations and bespoke image formats.

Q12 How many images are in your collection now?		
	Respons e Percent	**Respons e Total**

1-100	5.8%	4
101-500	4.3%	3
501-1000	8.7%	6
1001-10000	**36.2%**	**25**
10001-50000	27.5%	19
Other (please specify) *	17.4%	12
	Total Respondents	**69**

Discussion: The majority of collections were in the range 1,001 to 10,000 images. It would be difficult to specify the total number of images stored across the respondents' image collections but a conservative estimate using the lowest number for each category and the absolute numbers given in the free text entry “other” column for the largest collections would suggest that respondents to this survey represent somewhere around 1.5 million images.

Q13 Do you consider the size of your collection to be static, or will it continue to grow?		
	Response Percent	**Response Total**
Static	17.1%	12
Growing	**82.9%**	**58**
	Total Respondents	**70**

Discussion: As expected the majority of the collections 82% considered themselves to be actively expanding their collections, with only 17% considering their collection to be static. Later questions suggest that often the reason for the collection to not take on any more material is that it has reached the end of an externally funded digitisation project and lack of funding or time prevents the collection taking on more images. Often the original development team has been disbanded at the end of the project.

Q14 How many images do you expect to reach in two years		
	Response Percent	**Response Total**
1-100	0	0
101-500	4.8%	3
501-1000	6.5%	4
1001-10000	27.4%	17
10001-50000	**33.9%**	**21**
	Total Respondents	**62**

Discussion: With 82% considering themselves still growing, and the category of 10,001-50,000 now moves to being the most popular category for the size of collections in two years time.

Q15 How many users does your collection have now?		
	Response Percent	**Response Total**
1-100	3.1%	2
101-500	3.1%	2
501-1000	9.4%	6
1001-10000	15.6%	10
10001-50000	21.9%	14
Other (please specify) *	**46.9%**	**30**
Total Respondents		**64**

Discussion: To an extent this is a difficult question to answer for any open access web based collection. The numbers of users can often only be estimated from web access logs and these are notoriously difficult to interpret, this was shown by nearly half the respondents to the question submitting an open ended free text explanation giving snapshot figures for visitors per day, or per year, or expressing the difficulty of calculating accurate figures. It appears that the more popular sites receive thousands of distinct visits per month. See Table 24 for how the collections log access statistics.

Q17 Do people outside your own institution/group/hobby contribute to the collection?		
	Response Percent	**Response Total**
Yes	**62.2%**	**43**
No	36.8%	25
Total Respondents		**68**

Discussion: A somewhat surprising 62% have people outside the immediate organisation contributing material to the collection. When questioned further on the process for submission to the collection, the most popular answer was simple mechanisms such as accepting slides and photos for digitisation later. Formal online submission was not mentioned.

Access

Q19 What access restrictions are there on your collection?		
	Response Percent	**Response Total**
Open access	**82.9%**	**58**
Only I can use it	0%	0
Restricted to users from your institution	1.4%	1
Restricted to educational users by ATHENs or other password system	2.9%	2

Subscription based (username/password protected) and free of charge	4.3%	3
Subscription based (username/password protected) and charge for use	2.9%	2
Other (please specify) *	24.3%	17
Total Respondents		**70**

Discussion: Nearly all the collections surveyed had an online web presence so it isn't surprising that 83% are open access. However it is interesting that very few implement subscription or password-based access control.

Rights

Q20 What licence restrictions do you impose on the use (download, printing, copying) of your material?		
	Response Percent	**Response Total**
Free access	38.6%	27
Non-commercial use only	30%	21
Educational use only	22.9%	16
Schools use only	7.1%	5
Use by institution only	5.7%	4
Personal use only	15.7%	11
Subscribers only	4.3%	3
Other (please specify) *	**45.7%**	**32**
Total Respondents		**70**
*** Other Responses**		

Discussion: As over 80% of the collections were open access, this question tries to refine what the user can do with the material viewed. This was a wide open question and response varied from 38% responding "Free Access" and 30% responding "non-commercial use", to 16% answering "personal use only". It is striking that most respondents (46%) felt compelled to give a free text response as the possible survey choices were too simplistic. Fifteen of those who gave a free text response stated that they would expect payment for use of their material.

A few quotes typify the range of attitudes:

> *"It is accepted that material from the database will be printed out for research purposes, or inclusion in student essays. All other use is governed by copyright, and permission must be sought."*
>
> *"Should be for personal use or educational use. We sell high quality photographic copies of images through the site and charge a reproduction fee for any commercial use."*
>
> *"No downloading, printing or copying permitted"*

A common policy was to restrict images on the web to low quality images or small pixel dimensions as a deterrent to misuse:

> *"Images on web are entered in low resolution, so that they cannot be downloaded in a form which could be re-used in publications"*

"Material can be supplied for research or private study, there is on-line access to low resolution images for subscribers. Any reproduction is only with written permission and license fee".

Q21 What conditions do you impose on the **re-use** (for web or print publication) of material from your collection? (please tick all that apply)		
	Response per cent	**Response total**
I do not allow re-use of material	11.4%	8
Attribution of you or your collection as a source	**55.7%**	**39**
Non-commercial re-use of your material only	30%	21
No modification of your material	20%	14
Modifications/derivatives of your work must be distributed under the same licence terms	5.7%	4
Geographical scope of your licence (e.g. limited to UK)	1.4%	1
Other (please specify) *	47.1%	33
Total respondents		**70**

Discussion: This question compares typical licence restrictions with image providers own terms and conditions for reuse of material on the web or in print. We based these categories on the Creative Commons licence types in the hope that we could match image provider's attitudes to CC licences.

12% of respondents state that no re-use of the material is possible. 30% stated "non-commercial use only" whilst 56% stated that attribution was always needed. Nearly 50% of the respondents needed to add other free text to clarify the terms and conditions. A typical example explaining "non-commercial use" was:

"permission to download, print and repurpose content from the site for educational purposes, providing that you do not sell or redistribute that material."

Another provider commented on commercial use:

"A reproduction rights fee will be negotiated for any re-use of an image for web or print publication."

Comments provided by respondents suggested that 18 collections would allow commercial use of their images on payment of a fee.

It is also interesting to realise that some providers are not themselves able to make the decision on reuse (6 respondents). In the words of an image provider:

" the work is copyright and permission for re-use should be requested of the owner of the work."

Q22 Do you use any form of textual or other watermarking to protect your material?		
	Response Percent	**Response Total**
None	**67.1%**	**47**
Textual/Graphical watermark	22.9%	16
Digital Watermarking	12.9%	9

Total Respondents	**70**

Discussion: This was a simpler question for the image providers to answer. Do you protect your material by using a watermark? The majority, 66% said no. A smaller percentage, 23% said that they added a textual or graphical icon stamp to their images. Surprisingly only 9 image providers, typically mature large collections with subscriptions aspects used Digital Watermarking. The commercial watermark software **Digimarc** (http://www.digimarc.com/) was mentioned by one of the respondents.

Q23	Does your collection have a declaration of copyright ownership, stating who owns the copyright?	
	Response Percent	**Response Total**
Yes	**91.4%**	**64**
No	8.6%	6
Total Respondents		**70**

Q24	Does the copyright declaration include contact information?	
	Response Percent	**Response Total**
Yes	**80.9%**	**55**
No	19.1%	13
Total Respondents		**68**

Q25	Does your collection have a formal statement of conditions of use?	
	Response Percent	**Response Total**
Yes	**75.7%**	**53**
No	24.3%	17
Total Respondents		**70**

Q26	Do you have a policy on allowing the use of the material in an educational context?	
	Response Percent	**Response Total**
Yes	**66.2%**	**45**
No	33.8%	23
Total Respondents		**68**

Discussion: Considering the four questions concerning the copyright declaration, contact information, conditions of use and educational context of use, we see a that the majority 92% have a declaration of copyright ownership on their collection, slightly fewer, 81% have contact information (which is important if you wish to clarify reuse terms and conditions), 76% have a statement on conditions of use and then 66% have a policy on educational use.

It is the latter question that is of particular interest as these image providers were all selected to be of interest to the educational community, yet one third do not have a defined policy that separates educational use from general public use. This question demonstrates the need for a national policy initiative for licences and terms and conditions.

Q27 What are your primary concerns about sharing the images in your collection with other people?	
Total Respondents	**43**

Discussion: Most respondents felt that sharing images more widely would result in unauthorised commercial reuse, and loss of copyright control. The most common concerns raised were as follows:

	No. Respondents
Unconcerned	5
Ethical concerns	7
Lack of attribution	8
Unauthorised commercial reuse	21
Loss of copyright control	25
Don't own copyright	5

For some providers (5 respondents) there were very few concerns, adopting the attitude that the purpose of the site was to provide greater access to the material:

> *"My primary concern is to make the images available to as many people as possible."*
>
> *"As this is an important religious and cultural manuscript, there are no real concerns about sharing it with its user community."*

The vast majority of concerns were about copyright; 25 responses mention it explicitly, and it is clear that many collections do not own the copyright to the images that they are displaying.

Twenty-one responses express concern about unauthorised commercial reuse if the collection were to be shared. Eight respondents mentioned lack of attribution. Seven respondents raised ethical concerns.

Another provider represented those that had taken a pragmatic approach of only offering low quality images online:

> *"Only low resolution (approximately 60dpi) images are offered on the web pages and these would not be suitable for commercial use. We are not concerned that images at low resolution can be downloaded. We would be concerned if it were possible to use the digital images in any commercial manner. The purpose of the digital image collection is by its very nature to share the pictures of the original collections with other people."*

Only one image provider specified a formal licence, and interestingly it is a Creative Commons (http://creativecommons.org) licence that users might already be aware of through social file sharing:

> *"That users adhere to the terms of use a Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License."*

Contextual information

Introduction: Two open-ended questions asked about any contextual information that was stored with the image. We asked if a formal metadata schema was used to catalogue the image. Dublin Core was the only frequently mentioned formal schema for metadata with many providers explaining that their own internal database fields could be mapped to Dublin Core.

	No. Respondents
No Schema used	23
Dublin Core	30
Own Schema	10
IPTC	2
Spectrum	2
VRA	2
Other (HURIDOCS, MARC, METS, NADS, PREMIS, RLG, TEI)	7

Q30 Does the collection allow users to annotate the material via comments etc?		
	Response Percent	**Response Total**
Yes	16.2%	11
No	**83.8%**	**57**
	Total Respondents	**68**

Discussion: Surprisingly few of the image providers, 16% have a method for users of the material to comment on the image through adding their own thoughts. This, coupled with the relatively few systems that allow users to see commonly viewed popular images (see Table 21) shows that there is quite a considerable difference between these educational image providers and their more commercially orientated stock photography and image sharing rivals. Presumably this lack of social feedback about the material is because of the extra layer of technical complexity and the option not being considered necessary in the early specification stage.

Q31 Which of the following options are available for searching and browsing your collection?			
	Currently available	**Would like to implement**	**Respondent Total**
Thumbnail display of browse or search results	**89% (50)**	11% (6)	**56**
Browse by predefined categories	**79% (46)**	21% (12)	**58**
Search using structured/controlled vocabularies	**78% (35)**	22% (10)	**45**
Search by free text	**76% (44)**	24% (14)	**58**
Search by type of original e.g. photograph	**72% (23)**	28% (9)	**32**

Download individual images and their associated catalogue information to the local PC	**69% (16)**	31% (6)	**35**
Search powered by Google or other external service	**63% (15)**	38% (9)	**24**
Ability to narrow and broaden searches after an initial request	38% (15)	**62% (25)**	**40**
Search using visual characteristics of the image (colour, layout, shape etc)	29% (5)	**71% (12)**	**17**
Browse by commonly-viewed images	24% (4)	**76% (13)**	**17**
		Total Respondents	**68**

Discussion: The most common way of browsing the providers images was by thumbnail display of browse or search results 89%, with pre-determined categories 79%, controlled vocabularies 78% and free text search 76%. Only 5 image providers had a mechanism for searching using visual characteristics of the image, though 12 collections said they would like to implement this option. Only 4 collections allowed users to browse commonly-viewed images and this was the most popular feature that collections would like to implement.

Q32 Do you currently log the activity and usage of your collection?		
	Response Percent	**Response Total**
No	11.6%	8
Server log files of individual downloads	**42%**	**29**
Monitoring of logins and user activity	**40.6%**	**28**
Other (please specify) *	27.5%	19
	Total Respondents	**69**

Discussion: Monitoring usage of an online collection is perhaps more difficult than at first seen particularly if the collection doesn't have a sophisticated dedicated back end image management system. The responses to this question ranged from "Full photographic library management system recording all rights management, financial and licensing and other usage of the collection." to "User stats supplied by ISP". The most common form of monitoring 42%, was analysis of the server log files of individual file downloads. Many providers mentioned professional packages, such as "Web trends" that were used to analyse the web server log files. These packages provide human readable web access statistics and trends in usage across a web site. However web site traffic analysis is a notoriously difficult task that is more suited to general trends across time in usage rather than absolute figures for access. A similar amount, 41% monitored logins and user activity.

Q34 Below are three hypothetical national services that could allow people to access individual collections of images. Through which ones (if any) would you allow people to access your collection?				
	Yes	**No**	**Maybe**	**Response Total**
A directory of digital collections	**84% (57)**	3% (2)	13% (9)	**68**
A central web site that holds information about the	**76% (51)**	7% (5)	16% (11)	**67**

contents of your collection				
A central repository of educational images	31% (21)	27% (18)	**45% (28)**	**67**
			Total Respondents	**68**

Discussion: A significant proportion of respondents would be interested in a directory of digital collections, presumably because there is no large effort or cost associated with buy-in. A similar number would be interested in a site that holds information about collection contents. This idea also sits comfortably with the commonly expressed worry that collections may get marginalised by large systems.

When asked later in the survey what services could be provided nationally (see Table 29) the second choice was Marketing and the third choice was search facilities. This does seem to add weight to the proposal for more work done on directories of image providers held centrally and then related marketing could be more easily done on behalf of these providers.

It is interesting that collection providers are less convinced about the creation of a central repository for images. Only 31% were supportive of allowing their collection to be part of a central educational repository. It is possible that this is also tied to the fear of marginalisation. One provider explained the pressures involved:

> *"We are keen to enhance access to our collection; this is a key strategic objective. However, our main difficulties/limitations relate to limited institutional resources for digitisation and addressing this is a much higher priority than delegating/transferring image activities to a national centre. That said, in principle we would support the development of network of image collections and associated support."*

Q35 How do you store your 'original' images and other files?		
	Response Percent	**Response Total**
Folders on a PC	32.4%	22
CDs or other removable media	61.8%	42
Stored on a server	**69.1%**	**47**
Other (please specify) *	19.1%	13
	Total Respondents	**68**

Discussion: Most collections (69%) were storing the original images and files on an internal server and also 62% were backing the original source materials up onto removable media such as CDs and DVD-R.

Q36 How do you manage your 'original' images and other files?		
	Response Percent	**Response Total**
Database developed in-house	**52.3%**	**34**
Image management software developed in-house	7.7%	5
Commercial image management software	27.7%	18
Free or Open source package	4.6%	3

Ad hoc or other (please specify) *	24.6%	16
	Total Respondents	**65**

Discussion: A surprisingly large number of image providers had resorted to developing a database in-house to manage their 'original' images presumably because commercial systems did not meet their needs in terms of cataloguing. Open source or free solutions for managing the original files were only used by 5% of providers.

When questioned further, the most popular software named was Microsoft Access (10) collections, database connectivity was often SQL based such as a MySQL. However there was a broad range of solutions used, from simple folder hierarchies on the server, through Filemaker Pro and Access databases to commercial systems such as iBase (http://www.ibase.com/) and one mention of a freely available open source institutional repository software Dspace (http://www.dspace.org/).

Technical information

Q38 How do you make your collection available to others?		
	Response Percent	**Response Total**
Held on a local PC (Offline)	11.9%	8
Internal network (Intranet)	17.9%	12
Commercial ISP's web server (Internet)	25.4%	17
In-house web server (Internet)	**62.7%**	**42**
Other (please specify) *	14.9%	10
	Total Respondents	**42**

Q39 What type of software do you use to make your collection available to others?		
	Response Percent	**Response Total**
Database developed in-house	**52.4%**	**33**
Catalogue software developed in-house	7.9%	5
Commercial image management package	28.6%	18
Free or Open Source catalogue software	7.9%	5
Ad hoc or other system (please specify) *	28.6%	18
	Total Respondents	**41**

Discussion: Most of the collections are delivered from an in house web server to the internet (63%) whilst a much smaller group use a commercial ISP web server (25%). The majority of collections surveyed are using a back-end database system developed in-house (53%) with a much smaller amount (29%) using a commercial package such a iBase (http://www.ibase.com/).

Q40 What obstacles did you encounter when setting up your image collection? (tick all that apply)		
	Response Percent	**Response Total**
Insufficient time	**55.7%**	**34**
Insufficient funding	54.1%	33
Own lack of technical knowledge	39.3%	24
Lack of contribution from colleagues	16.4%	10
Lack of IT support	34.4%	21
Uncertainty about copyright	29.5%	18
Lack of common cataloguing (metadata) standards	29.5%	18
Gathering up existing image collections held by staff	14.8%	9
Gaining access to physical objects to digitise	8.2%	5
Lack of suitable database systems for storing and making available the images	27.9%	17
Lack of archival/backup capabilities	27.9%	17
Other (please specify) *	24.6%	15
Total Respondents		**61**

Discussion: The three biggest hurdles were unsurprisingly lack of time, lack of funding and also lack of technical knowledge. A number of the providers had received money original from digitisation schemes such as the heritage based National Opportunities for Funding (NOF) and this had now finished often frustrating the image providers who had expected to catalogue and expose more material but the initial stages of technical infrastructure and IPR clearance had often taken much longer than expected.

Q43 What services would you like to be provided by a national support network for image collections?		
	Response Percent	**Response Total**
Technical Support	**45.3%**	**29**
Marketing	42.2%	27
Search facilities	40.6%	26
Backup/Archive	34.4%	22
Payments and royalty collection	29.7%	19
Hosting of the collection	29.7%	19
User management	21.9%	14
Software provision	15.6%	10
Other (please specify) *	32.8%	21
Total Respondents		**42**

Discussion: Technical support is the most requested form of service that could be offered centrally (45%) followed by Marketing then Search facilities. A national technical support service is an obvious solution supported by Q40 which mentioned that the third most popular hurdle to collection building after funding and lack of time was lack of technical knowledge.

A national search service was supported by 41% of the respondents, this perhaps might be a cross-search facility, a portal gateway to image collections or perhaps the collections themselves had limited search facilities due to lack of technical knowledge or infrastructure and this could be solved by a national service that had access to the sites catalogue information. Again this adds some support to the idea of a centrally located site that has knowledge of the catalogue contents of the individual collections.

11.4.2 Survey Questionnaire

The CLIC (Community Led Image Collections) study is reviewing the growth of community-based digital image collections, in the UK educational sector.

CLIC will make recommendations to the JISC Image Working Group on sharing and embedding collections within Higher and Further Education.

Community-Led Image Collections September 2005

Image collection providers questionnaire

Contact: Jonathan Miller, Learning Technologies Group, University of Oxford,

13 Banbury Road, Oxford, OX2 6NN

Web Site: http://clic.oucs.ox.ac.uk/ Email: clic@oucs.ox.ac.uk

Who should complete the questionnaire?

This questionnaire is for the maintainers or owners of digital image collections.

It contains 44 questions, and will take approximately 20 minutes to complete.

We will treat your responses to this questionnaire confidentially. Please see the confidentiality statement at the end of the questionnaire.

What do we want to find out?

What technical and infrastructure needs does your collection have? How could these be solved by a national support service for image collections?

What information do you hold about the images in your collection? How much of this would be meaningful to a national image-search directory?

Does your collection serve a particular community? Is there scope for sharing your collection more widely?

What is your attitude to sharing your images with others? Would you be happy to allow open access to your collection for educational purposes?

What are the key obstacles to the growth of your collection? What technologies would help overcome these obstacles?

Purpose of the questionnaire

The CLIC study is looking at the growth of community-led educational digital image collections. The scoping study will make recommendations to the JISC Image Working Group on possible software, infrastructure and network needs that could be satisfied at a national or local level.

Specifically, the study will:

- Survey the nature of current image collections.
- Review the obstacles encountered by providers during collection building.
- Make recommendations that could be implemented at a national level.

Collection Details

1. What is the title or name of your collection?

2. Name of contact person for the collection

3. Institution or group

4. Email address

5. Collection website

6. Please describe your involvement with this image collection (tick any that apply)

- ☐ I am the main contributor to the collection
- ☐ The collection is a major responsibility of my employment
- ☐ The collection is one of many responsibilities of my employment
- ☐ The collection is a sideline activity or hobby
- ☐ Other please specify_____

7. Please provide a short description of the collection and the subject areas it covers:

_____

8. Please indicate the subject(s) to which your collection is relevant:

- ☐ Medicine and Dentistry
- ☐ Nursing, Anatomy, Physiology and Pathology
- ☐ Biological Sciences, Biology, Zoology, and Psychology
- ☐ Vetinary Sciences, Agriculture, Forestry
- ☐ Physical Sciences, Chemistry, Materials Science, Physics and Geology
- ☐ Mathematical, Operational Research, Statistics and Computer Sciences
- ☐ Engineering
- ☐ Technologies, Metallurgy, Ceramics and Materials Technology
- ☐ Architecture, Building and Planning
- ☐ Social Studies, Economics, Politics, Anthropology, Human and Social Geography
- ☐ Law
- ☐ Business and Administrative Studies, Finance, Accounting, Marketing & Accounting
- ☐ Mass Communications and Documentation, Media Studies, Publishing and Journalism
- ☐ Linguistics, Classics, English studies and related subjects
- ☐ European Languages, Literature and related subjects
- ☐ Eastern, Asiatic, African, American and Australasian Languages and Literature
- ☐ Historical, Archaeology and Philosophical studies
- ☐ Creative Arts, Music, Cinema, Photography and Design
- ☐ Education
- ☐ Other please specify_____
  _____
  _____

9. Please indicate the age range(s) to which your collection is relevant:

- ☐ Primary Education (4-11)
- ☐ Secondary Education (11-16)
- ☐ Higher Education (University)
- ☐ Life-long learners/General public (Adults)

☐ Further Education (16-21)

☐ Other please specify_____

10. What are the sources of the digital images in your collection?

☐ Existing collection of slides/photos

☐ Existing collection of digital images

☐ Photos/slides created specifically for the collection

☐ Digital photos/scans created specifically for the collection

☐ Other please specify_____

11. If your collection contains any other types of media files, please indicate what these are:

☐ Audio files (MP3, RealAudio, etc.)

☐ Video files (MPEG, WMV, QuickTime, etc.)

☐ 3D modelling files (VRML, QuickTime VR, etc.)

☐ Other please specify_____

12. How many images are in your collection now?

☐ 1-100 ☐ 101-500 ☐ 501-1,000 ☐ 1,001-10,000 ☐ 10,001-50,000
☐ Other please specify_____

13. Do you consider the size of your collection to be static, or will it continue to grow?

Static ☐ Growing ☐

14. How many images do you expect to reach in two years?

☐ 1-100 ☐ 101-500 ☐ 501-1,000 ☐ 1,001-10,000 ☐ 10,001-50,000
☐ Other please specify_____

15. How many users does your collection have currently?

☐ 1-100 ☐ 101-500 ☐ 501-1,000 ☐ 1,001-10,000 ☐ 10,001-50,000
☐ Other please specify_____

16. How many people have contributed, or will contribute images to the collection?

Please specify

17. Do people outside your own institution/group/hobby contribute to the collection?

Yes ☐ No ☐

18. If yes, what is the process for submitting material to the collection? (e.g. they create a jpeg file which is emailed to you for review)

Please specify

Access and rights

19. What access restrictions are there on your collection?

- ☐ Open access
- ☐ Only I can use it
- ☐ Restricted to users from your institution
- ☐ Restricted to educational users by ATHENs or other password system
- ☐ Subscription based (username/password protected) and free of charge
- ☐ Subscription based (username/password protected) and charge for use
- ☐ Other please specify _____

20. What licence restrictions do you impose on the use (download, printing, copying) of your material?

- ☐ Free access
- ☐ Educational use only
- ☐ Schools use only
- ☐ Use by institution only
- ☐ Subscribers only
- ☐ Personal use only
- ☐ Non-commercial use only
- ☐ Other please specify _____

21. What conditions do you impose on the re-use (for web or print publication) of material from your collection? (please tick all that apply)

- ☐ I do not allow re-use of material
- ☐ Attribution of you or your collection as a source
- ☐ Non-commercial re-use of your material only
- ☐ No modification of your material
- ☐ Modifications/derivatives of your work must be distributed under the same licence terms
- ☐ Geographical scope of your licence (e.g. limited to UK)
- ☐ Other please specify _____

22. Do you use any form of textual or other watermarking to protect your material?

None ☐ Textual/Graphical watermark ☐ Digital watermarking ☐

23. Does your collection have a declaration of copyright ownership, stating who owns the copyright?

Yes ☐ No ☐

24. Does the copyright declaration include contact information?

Yes ☐ No ☐

25. Does your collection have a formal statement of conditions of use?

Yes ☐ No ☐

26. Do you have a policy on allowing the use of the material in an educational context?

Yes ☐ No ☐

27. What are your primary concerns about sharing the images in your collection with other people?

Please specify

Contextual information

28. What subject-specific information do you store with the image?

Please specify

29. Do you use a formal metadata schema (IPTC, Dublin Core, METS, VRA3) to catalogue your material?

Please specify

30. Does the collection allow users to annotate the material via comments etc?

Yes ☐ No ☐

Searching and browsing for images

31. Which of the following options are available for your collection?

Currently available	Would like to implement	
☐	☐	Search by free text
☐	☐	Search powered by Google or other external service
☐	☐	Search using structured/controlled vocabularies
☐	☐	Search by type of original e.g. photograph
☐	☐	Search using visual characteristics of the image (colour, layout, shape etc)
☐	☐	Ability to narrow and broaden searches after an initial request
☐	☐	Browse by predefined categories
☐	☐	Browse by commonly-viewed images
☐	☐	Thumbnail display of browse or search results
☐	☐	Download individual images and their associated catalogue information to the local PC

32. Do you currently log the activity and usage of your collection?

☐ No

☐ Server log files of individual downloads

☐ Monitoring of logins and user activity

☐ Other please specify _____

33. Do you see any value in having the contents of your collection revealed in search engines such as Google Image Search?

Please specify

34. Below are three hypothetical national services that could allow people to access individual collections of images. Through which ones (if any) would you allow people to access your collection?

	Yes	No	Maybe
A directory of digital collections	☐	☐	☐
A central web site that holds information about the contents of your collection	☐	☐	☐
A central repository of educational images	☐	☐	☐

Technical aspects

35. How do you store your 'original' images and other files?

- ☐ Folders on a PC
- ☐ CDs or other removable media
- ☐ Stored on a server
- ☐ Other please specify_____

36. How do you manage your 'original' images and other files?

	Name of software/database/development technology?
☐ Database developed in-house	_____
☐ Image management software developed in-house	_____
☐ Commercial image management software	_____
☐ Free or Open source package	_____
☐ Ad hoc or other system	_____

37. How do you archive your 'original' image files?

Please specify

38. How do you make your collection available to others?

- ☐ Held on a local PC (Offline)
- ☐ Internal network (Intranet)
- ☐ Commercial ISP's web server (Internet)
- ☐ In-house web server (Internet)
- ☐ Other please specify_____

39. What software do you use to make your collection available to others? Please give details.

Name of software/database/development technology

- ☐ Database developed in-house _____
- ☐ Catalogue software developed in-house _____
- ☐ Commercial image management package _____
- ☐ Free or Open Source catalogue software _____
- ☐ Ad hoc or other system _____

40. What obstacles did you encounter when setting up your image collection? (tick all that apply)

- ☐ Insufficient time
- ☐ Insufficient funding
- ☐ Own lack of technical knowledge
- ☐ Lack of contribution from colleagues
- ☐ Lack of IT support
- ☐ Uncertainty about copyright
- ☐ Lack of common cataloguing (metadata) standards
- ☐ Gathering up existing image collections held by staff
- ☐ Gaining access to physical objects to digitise
- ☐ Lack of suitable database systems for storing and making available the images
- ☐ Lack of archival/backup capabilities
- ☐ Other please specify…_____

41. Which were the 3 biggest of these obstacles, and how were they addressed?

Please specify

42. What are the biggest obstacles to further developing the collection?

Please specify

43. What services would you like to be provided by a national support network for image collections?

- ☐ Backup/Archive
- ☐ Technical Support
- ☐ Hosting of the collection

- ☐ Search facilities
- ☐ User management
- ☐ Marketing
- ☐ Payments and royalty collection
- ☐ Software provision
- ☐ Other, please specify_____

44. We welcome any further comments on the needs of your collection, or any ideas that will help inform the CLIC project…

Thank you for taking the time to complete this questionnaire.

Please return the completed questionnaire to
Jonathan Miller
Learning Technologies Group,
University of Oxford,
13 Banbury Road,
Oxford,
OX2 6NN

Follow up and confidentiality

1. Follow-up interview: ☐ Please tick if you would be happy to be contacted by email with further questions.

2. Survey findings: ☐ Please tick if you would like to receive our summary of the survey results.

3. Confidentiality

Your information will be kept confidential, and used anonymously. We will not share the information you provide unless you give us permission.

The CLIC project is working with a number of other investigations, also funded by JISC, and it may be that information you provide would assist these other projects.

If you give us permission for some or all of the information to be shared with other JISC projects, please complete this table:

- ☐ I give permission for you to share the entire questionnaire
- ☐ I give permission for you to share the entire questionnaire, but with the following conditions

  _____
  ___

11.5 WP5 Survey of Repository Software for Image Management

The following table reviews 18 pieces of software that can be used to manage images.

Name	4Images Gallery	ADLIB (ADLIB Archive/Library/Museum)	Canto Cumulus	CopperMine
Project Information	Website: http://www.4homepages.de/ (Mainly German, some English) Company: Dots United.	Website: http://www.adlibsoft.com/ Company: Adlib Information Systems	Website: http://www.canto.com/ Company: Canto Software, Inc.	Websites: http://coppermine-gallery.net/ http://sourceforge.net/projects/coppermine/ Project Managers: Joachim Müller, Dr Tarique Sani.
Requirements	• Web server (e.g. Apache) • PHP (> v4.0.5) • SQL Database: MySQL (> v3.23) • Image manipulation package: GD Lib, ImageMagick or NetPBM. (Optional)	• Windows Web server (e.g. Microsoft IIS) (Note that for the ability to browse the catalogue over the internet, "Adlib Internet Server" is required, as well as the appropriate version of Adlib)	• Mac, Windows or UNIX PC. • Appropriate server, for users' requirements.	• Web server (e.g. Apache, Microsoft ISS) • PHP • SQL Database: MySQL. • Image manipulation package: ImageMagick or GD Lib.
Technology	• PHP/MySQL. • OS Independent.	• OS: Windows	• Java (Embedded Java Plugins) • Windows or Mac OS X (or UNIX for the Server)	• PHP/MySQL. • OS Independent.
Installation	Upload to web server required, followed by web based configuration.	Installer provided. A windows web server is required if the catalogue needs to be accessed across a network/the internet.	Installer provided for both client and server systems	Relatively straight forward, installer is provided, and also straight forward configuration.

User Interface	• Web based. • Appearance editable by changing HTML templates.	• Web based for viewing the catalogue over a network. • Window based for viewing/managing the catalogue on the main server.	• Uses Cumulus Client software (to access Server software)	• Web based. • Appearance customisable using web interface, CSS, PHP or plain HTML.
Ease of Use	Some basic web server and file system knowledge to install, although does include installer. Simple web-based user interface for browsing. Images uploaded by admin only, using web-based interface, or by FTP.	Some basic web server and file system knowledge required to install (if allowing access over a network). Web-based searching for users. Uploading done by a standard Windows interface.	Installer provided for both client and server systems. May require some basic technical knowledge to get them all working together. Provides simple windowed interface for using the software.	Some basic web server knowledge required to install. Simple web-based user interface for browsing. Pictures can also be uploaded by simple web interface, or by FTP.
Sharing	User based access to gallery. Users have different levels of access and for different galleries, and can upload images depending on their access level.	Users over a network have access to browsing and searching the gallery. Images uploaded only on the main server.	Client systems have user based access to the main server gallery. Users can access, view, and upload files to the main gallery.	Anyone can browse the gallery (using the web-based interface) and use the images. Some albums can be made private and require a username and password to access. Access to private galleries can be setup for individual users/groups of users.

Metadata	No standards implemented. Only contains optional description field and keywords, but can be nested within user-defined categories.	ISAD(G) 2nd edition standard implemented (http://www.ica.org/biblio/cds/isad_g_2e.pdf). Allows comprehensive searching of metadata, as well as custom indexing.	Allows large numbers of custom fields to be added on a per-category basis, including defining which particular fields should be included when searching. Data can be added once images have been uploaded	No standards implemented. Contains title and description, but does allow user defined fields. These fields are searchable. Details are submitted in a form.
License	May be used and modified free of charge for personal and non-profit use. Commercial use requires purchase of a licence. €99/£67 licence for private/commercial use.	Can be purchased for commercial or large scale use. The source for the software is not available.	Can be purchased for personal use, or large scale use for varying prices. Source for the product is not available.	GNU General Public License. Open source, free to use and modify.
Rights	No rights management	No rights management.	No rights management.	No rights management.
Features	• Images are stored in nested categories. • Allows user comments and rating for individual images. • Allows numerous file formats including audio and video.	• Emphasis seems to be on searching for the required images, as opposed to creating albums etc.	• Fully customisable with the Embedded Java Plugins (and some java knowledge). • Supports use of EXIF data.	• Images are arranged in "categories" and "albums". • Allows private galleries. • User management integration with various BBSes. • Supports use of EXIF data.
Name	DAlbum	Exhibit Engine	Gallery	Greenstone Digital Library Software
Project Informati	Websites: http://www.dal	Website: http://www.photogr	Websites: http://gallery.menalt	Website: http://www.gr

on	bum.org/ http://sourcefo rge.net/project s/dalbum/	aphy-on-the.net/ee/	o.com/ http://sourceforge.n et/projects/gallery/ Project Managers: Robert Balousek, Bharat Mediratta.	eenstone.org/
Require ments	• Web server (e.g. Apache) • PHP (> v4.1) • Image manipulation package: ImageMagick, NetPBM or GD Lib (Optional).	• Web server (e.g. Apache, Microsoft ISS) • PHP • SQL Database: MySQL	• Web server (e.g. Apache, Microsoft ISS) • PHP • SQL Database: MySQL, Oracle or PostgreSQL (Only for v2, and not v1) • Image manipulation package: ImageMagick, GD Lib, NetPBM or GraphicsMagick.	• Web server (e.g. Apache) (Not necessarily required, see Installation section) • Java
Technol ogy	• PHP. • OS Independent.	• PHP/MySQL. • OS Independent.	• PHP/MySQL. • OS Independent.	• Mainly written in C++ • OS: Any Windows or UNIX
Installati on	Basic web server installation.	Some web server and file system based steps to install.	Straight forward copy to web root. Simple web configuration.	Installer for main software. Then some web server and java setup required. If, however a "local library" setup is used, no separate web server is required (only available on Windows).
User Interface	• Web based. • Appearance customisable by using PHP, CSS or HTML templates.	• Web based. • A number of options can be changed to customise the layout, and CSS can be used.	• Web based. • Appearance customisable using built in themes, CSS, or plain HTML, (but only allows a grid format).	• Web based. • Customisable , but only with some in-depth knowledge of the system (there are guides available to

				customisation).
Ease of Use	Some basic web server knowledge required to install. Simple web-based user interface for browsing and management. Image uploading is done solely by FTP (although separate PHP file managers are compatible).	Some basic web server knowledge required to install. Web-based user interface for browsing. Uploading of images is by FTP.	Some basic web server and file system knowledge required to install. Simple web-based user interface for browsing. Pictures can be uploaded by web interface or by FTP.	Setup requires some web server knowledge, and possibly a small amount of java knowledge. Simple web-interface for browsing and uploading.
Sharing	Anyone can browse the gallery. Allows password protected albums.	Anyone can browse the galleries, but passwords can be set on required galleries.	User based access to the gallery. Users can be given their own albums to store their own images. Access to albums is dependant on the user's access privileges, and has many different levels of access including just viewing, editing details, editing image, etc.	User based access, to determine whether a user can access or perform admin operations on the library. Allows full searching of metadata. Images/Files can be stored across a number of different machines/servers and still be shown in the same collections.
Metadata	No standards implemented. No extra metadata is stored.	No standards implemented. Completely customisable fields can be assigned to the galleries. Fields are always optional, and are displayed together	No standards implemented. Contains bare minimum of title and description.	Dublin Core metadata standard is implemented. Completely configurable to add new metadata sets.

		with each image.		Metadata can be exported as XML.
License	GNU General Public License. Open source, free to use and modify.	Unknown	GNU General Public License. Open source, free to use and modify.	GNU General Public License. Open source, free to use and modify.
Rights	No rights management.	No rights management.	No rights management.	No rights management.
Features	• Windows Explorer style interface. • EXIF data shown if found (optional). • Support for images which should not be resized (e.g. maps etc). • Bare minimum (if any) configuration required. • Titles/comments provided using plain text/html files.	• Images can be on any server - distributed gallery load possible. • Each image can have any amount of variations (size, colours, sharpening whatever). • Hidden galleries (password protected). • Hidden photos - keep some photos out of public view. • Administrator editors are all GUI and very powerful.	• Images are arranged in "galleries" which themselves contain (nested) "albums", which can be seen as a tree structure. • Supports use of EXIF data. • Public commenting. • Gallery publishing in RSS. • Individual albums have customisable appearance. • Allows private galleries/albums.	• Plugins (either existing, or user created) can be used to allow more different file types to be used (e.g. HTML, Word, etc). • Designed for large (multi-Gb) collections. • Collections can be published to the internet or onto CD-ROM.
Name	iView MediaPro (iView Multimedia)	LinPHA ("Linux Photo Archive")	MDID (Madison Digital Image Database)	MediaDB
Project Information	Website: http://www.iview-multimedia.com/	Websites: http://linpha.sourceforge.net/ http://sourceforge.net/projects/linpha	Websites: http://www.mdid.org/ http://sourceforge.net/projects/mdid	Website: http://et.middlebury.edu/et/new/software/mediadb/
Requirements	• Mac or Windows PC. • Some (undemanding) PC requirements.	• Web server (e.g. Apache) • PHP (> v4.1) • SQL Database: MySQL or Postgres • Image	• Web server (e.g. Microsoft IIS on Windows, or Apache on Linux) • SQL Database: MySQL or Microsoft SQL Server	• Web server (e.g. Apache) • PHP (v4) • SQL Database: MySQL • Image

		manipulation package: ImageMagick or GD Lib.	• Microsoft .NET	manipulation package: ImageMagick
Technology	• OS: Mac OS 9 or OS X, or Windows	• PHP/MySQL. • JavaScript. • OS Independent.	• Microsoft .NET (C# and ASP.NET) • JavaScript • OS: Windows and Linux (using "mono")	• PHP/MySQL. • OS Independent.
Installation	Full installer provided.	Web-based installation. Web configuration.	Project compiling (using Nant). File management and web server configuration steps.	Extraction to web server, and some PHP configuration needed.
User Interface	• Single window • Tree type structure for storing images/files, which are shown separately or in a grid type view.	• Web based. • Appearance editable with CSS.	• Web based.	• Web based. • Limited appearance customisability.
Ease of Use	Standard windowed program interface, allowing simple navigation between images/files. Images easily added via menus/windows or drag and drop.	Some basic web server and file system knowledge required to install. Simple web-based user interface for browsing. Pictures uploaded by a web form, only for admin.	Some file system and server knowledge required to install successfully (a large number of tricky configuration steps). Simple web-based user interface for browsing.	Some PHP knowledge required to install, as well as web server configuration. Simple web-based user interface for browsing individual collections.
Sharing	Single system software. Users can produce web pages from their galleries/albums.	Anyone can browse the gallery, using the web-based interface. Admin is allowed to upload/manage images. Image information is added/edited after upload.	User based access to the catalogue including appropriate privilege management, although "guests" can be allowed. Images are generally searched for, and when loaded shows all the	Anyone can browse the gallery, but to access certain areas a login is required.

			metadata including copyright message.	
Metadata	Some (minor) standards implemented, including IPTC/ANPA, XNP (Adobe), and more. Allows export to XML files, or plain text tables.	No standards implemented. Contains bare minimum of tile and description. Allows keyword searching.	No specific standards implemented. Many custom fields can be entered. Fields can be used more than once (useful for searching the "subject" field for example).	No standards implemented. Very limited information can be associated with each image/collection.
License	iView End User License Agreement, allows single user use of product. Source/commercial use for the product is not available.	GNU General Public License. Open source, free to use and modify.	GNU General Public License. Open source, free to use and modify.	Unknown
Rights	No rights management.	No rights management.	Copyright details are added to the metadata, and shown whenever an image is opened up to be shown, whilst browsing.	No rights management.
Features	• Allows a huge variety of different image, and other formats. • Supports use of EXIF data. • Allows import from CDs/DVDs, URLs, and many more sources.	• Supports video files. • Permissions set for individual albums. • Supports use of EXIF data. • Download images in zip/tar/etc archive.	• Support for multiple collections. • Custom catalogue data structures. • Cross-collection searching. • User image notes and annotations. • Web-based slideshow viewer. • Image Viewer classroom application. • Packaged slideshows for offline presentation. • Printable flashcards. • Data exchange through XML.	• Support for multiple collections • Search and browse functions including cross-collection searching • Tools for managing slideshows • Web-based slideshow viewer. • Zoom functionality • Tools for managing user accounts and authenticatio

				n
Name	MediaWiki	Mig (My Image Gallery)	nGallery	Open Searchable Image Catalogue
Project Information	Websites: http://www.mediawiki.org/ http://sourceforge.net/projects/wikipedia	Website: http://mig.sourceforge.com/ http://sourceforge.net/projects/mig/	Website: http://www.ngallery.org/	Websites: http://sourceforge.net/projects/osic-win http://osic-win.sourceforge.net/
Requirements	• Web server (e.g. Apache) • PHP (> v4.3) • SQL Database: MySQL (> 3.23)	• Web server (e.g. Apache, Microsoft IIS) • PHP (> v3.0.9) • Image manipulation package: ImageMagick (Optional). • Perl (Optional).	• Web server (Microsoft IIS) • SQL Database (Microsoft SQL Server) • Microsoft .NET	• Web server (e.g. Apache) • PHP • SQL Database: MySQL • Image manipulation package: GD Lib. • Java Runtime Environment (Only for uploading files).
Technology	• PHP/MySQL. • ispell/aspell (Optional for spell checking) • OS Independent.	• PHP • Perl (only for image manipulation) • OS Independent.	• Microsoft .NET (C# and ASP.NET) • OS: Windows and Linux (using "mono")	• PHP/MySQL. • OS Independent.
Installation	Web server installation. Web based configuration.	Some basic web server and file management steps to install.	Installer provided.	A number of file-management based steps.
User Interface	• Web based. • Appearance editable by editing PHP. Also includes skins and special extensions.	• Web based. • Appearance editable using HTML templates.	• Web based. • Limited customisation possibilities.	• Web based. • Appearance changeable (text and colours only) for admin only.
Ease of	Some basic web server,	Some basic web server and file	Full web based interface for normal	Some basic web server

Use	file system and MySQL knowledge required to install. Simple web-based user interface for browsing. Images/files can be uploaded by web interface.	system knowledge required to install. Simple web-based user interface for browsing. Images are uploaded using the web server's file system only.	users and admin. Simple installation, provided all the Microsoft server settings are correct.	and file system knowledge to install. Simple web-based interface for browsing. Images uploaded by web form, by admin only.
Sharing	Completely open access to the gallery, anyone can add their own contributions. Previous versions of files are kept, so that spam, etc. can be easily removed, and pages/files restored to their previous state. Allows some user-based access for uploading files/images, etc.	Open access to the gallery. Images can only be added from the server itself. It is possible to set username/password access to specific folders.	Anyone can access the galleries, although each gallery can be optionally password protected. Admin can upload/manage images.	Anyone can browse the gallery, using the web-based interface. Only admin can manage/upload images. Image information is added for each image as they are uploaded using a web form.
Metadata	No standards implemented. Allows users to associate any (limited HTML formatted) text on pages, and add in the appropriate images. Includes full text search.	No standards implemented. No extra data is stored.	No standards implemented. Only title and description fields can be associated with each image.	No standards implemented. User defined fields can be assigned for separate galleries. Default fields include date, source, notes and description. Data entered during image upload on a web form.
License	GNU General Public License.	Similar licence to GPL. Open source, free to use and	GOTDOTNET Workspaces Commercial	GNU General Public License.

	Open source, free to use and modify.	modify.	Derivatives License (http://www.gotdotnet.com/workspaces/license.aspx?id=201eb290-5bf0-4452-9bf7-d21d39268f36). Similar to GNU General Public License.	Open source, free to use and modify.
Rights	No rights management, although would allow copyright notices to be put (more-or-less) anywhere on a page.	No rights management.	No rights management.	No rights management.
Features	• Designed to handle a large number of users and pages without imposing a rigid structure. • Pages are linked using user-assigned links on other pages. • Full web-based page editing and file upload features fully in place.	• Nested folders allowed for different albums. • Simply uses file system to determine which images (files) are shown in which albums (folders).	• User comments system. • Uses a gallery system.	• Emphasis on large numbers of images and the ability to search. • Images organised into galleries which each have (nested) categories.
Name	Qdig	YaPig		Other Pieces of Software
	(Quick Digital Image Gallery)	(Yet Another PHP Image Gallery)		
Project Information	Websites: http://qdig.sourceforge.net/ http://sourceforge.net/projects/qdig	Websites: http://yapig.sourceforge.net/ http://sourceforge.net/projects/yapig	The following pieces of software have not been included in this report for the given reasons: • OpenTerracotta (http://sf.net/projects/terracotta). This project seems to be no longer supported, and there is	

			a lack of information regarding the project available on the internet. • PhotoPost PHP (http://www.photopost.com/). Software seems to be lacking, due to all effort being put into the "Pro" (non-open source) version.
Requirements	• Web server (e.g. Apache) • PHP	• Web server (e.g. Apache) • PHP • Image manipulation package: GD Lib.	
Technology	• PHP • OS Independent.	• PHP • OS Independent	
Installation	Copy to web server, and some file permission setting steps. Can copy image files in a folder structure across at same time to instantly have set of galleries.	Web server installation.	
User Interface	• Web based. • File upload can optionally be done just by copying files into file structure. • Appearance customisable using some (simple) PHP to define variables.	• Web based. • Appearance editable with CSS.	
Ease of Use	Some basic web server and file system knowledge to install. Simple web based interface for browsing. Images can be uploaded using FTP, or by simply adding new images into the appropriate directory.	Some basic web server and file system knowledge required to install. For admin to add a new gallery, also requires some basic file system manipulation (including a lot of permissions changing). Everything else is web based including the uploading of images.	
Sharing	Anyone can browse the gallery,	Anyone with a password can browse the gallery, although if a user knows	

	using the web based interface. Image upload is done solely by admin, and can be done in multiple ways (as outlined above). Files can be accessed by FTP server.	the filename of a required picture, they can reach it without the password! Only admin can manage/upload images.
Metadata	No standards implemented. Only a description field is used.	No metadata standards implemented. No information stored with the images.
License	GNU General Public License. Open source, free to use and modify.	GNU General Public License. Open source, free to use and modify.
Rights	No rights management.	No rights management.
Features	• Web layout involves a row of images along the bottom of the page which can be selected, to give a full thumbnail view of a particular gallery. • Multiple (nested) galleries can be used. • FTP access.	• Allows user comments (optionally).

11.6 WP6 User Needs Conference

Subject areas represented at the July conference:

Institution or Department	Subject Area
Edinburgh University Data Library	Many subjects / general.
Digital Library Oxford	English
Medical Sciences Division	Medicine
Modern History Faculty	Oxford's historical portraits
Archaeology HCA	Archaeology
Department of Archaeology & Anthropology	Archaeology
School of Medicine	Medical
Department of Archaeology & Anthropology	Archaeology & Anthropology
Learning Centre	Public Art / Tourism / Architecture
Institute of Archaeology	Archaeology
Library Services	Mainly art, design and architecture
Archaeology	Archaeology: Anglo-Saxon specialists
Institute of Archaeology	Archaeology
School of Nursing	General Learning Objects
Communication Studies	The paintings of Walter Spies
Curatorial	Archaeology of SE Wales
The Pitt Rivers Museum	Ethnography and archaeology
The Pitt Rivers Museum	Anthropology and archaeology
Faculty of Design	High Wycombe Furniture
Forced Migration Online, Refugee Studies Centre, Oxford	Forced migration (the movements of refugees)
Photography and Art	Art and design, other areas
Library	Art and architecture
School of Medicine	Medicine, Pathology
History of Art	Art and General
School of Film, Music and New Media	Born digital art
The Pitt Rivers Museum	Anthropology/Ethnography
Sir William Dunn School of Pathology	Medical research image output
Library & Archive	Art and Art History
Learning and Interpretation	
Centre for Applied Research in Educational	Firdausi's Shahnama

Technologies	(Book of Kings)
Media Workshop	Education
Institute of Archaeology	Jericho archaeology
University Museum of Natural History	Natural History
Centre for Computing in the Humanities	Classics
Department of Earth Sciences, University of Cambridge	Fossils, rocks and minerals
The Pitt Rivers Museum	Anthropology and Ethnography
Language Training	Languages and Culture
History of Art	History of Art & Architecture
Learning Scotland	Many subjects / secondary education.
Department of Earth Sciences	Earth Sciences
Department of Chemistry	Chemistry

11.7 WP7 Visualising Science Survey

See pdf at

http://clic.oucs.ox.ac.uk/docs/WP7VisualizationSurvey.pdf

11.8 The Digital Picture in Arts Survey: Community aspects Report to CLIC

11.8.1 Introduction

This report is based upon the extensive research and consultation exercise carried out by The Digital Picture project, commissioned by the JISC Images Working Group and run by AHDS Visual Arts from April to September 2005.[1] It calls upon the project's report,[2] in turn based upon:

- A national consultation of the affected community and all associated parties using a questionnaire
- The creation of a consortium to represent the community via seminars and online conferencing
- Expert seminars to discuss a number of specified issues
- A literature/peer project review
- A formal response from the Association of Art Historians

The Digital Picture questionnaire asked a series of simple questions on a range of subjects that were considered to be important, and also included blank areas for more subjective responses. 502 individuals from over 150 institutions responded, including forty universities and over thirty associated museums, galleries and heritage organisations. The respondents' profile is broken down further in table 1.

Role	%	Involvement in	%	Age	%
Student	15	Further education	17	Under 18	1
Support staff	8	Higher education	40	18 – 30	21
Artist	8	Masters	17	31 – 50	45
Lecturer	28	PhD study	13	Over 50	32
Managerial	9	Post doctoral	7	Blank	1
Researcher	10	Other	6		
Librarian	12				
Art historian	5				
Other	5				

Table 1: breakdown of respondents to The Digital Picture[3]

The expert seminars were held from April to July 2005 at:

- The University of St Andrews

[1] http://thedigitalpicture.ac.uk/

[2] *The Digital Picture: a future for digital images in UK arts education*, http://thedigitalpicture.ac.uk/documents/pdf/digital_picture_final_report.pdf (consulted 6 December 2005). All further references are to section (§) and page (p.) numbers in this document, unless otherwise stated.

[3] § 4.1 on p. 10.

- Glasgow School of Art
- The University of Northumbria at Newcastle
- The University of Ulster, Belfast
- The University of Manchester
- Aston University, Birmingham
- The University of Wales, Cardiff
- The Courtauld Institute of Art, London
- The University of Bristol
- Surrey Institute of Art and Design
- Plymouth School of Art and Design

They were attended by librarians, academics, lecturers, artists, art historians, learning support workers, managers and students. Much of the discussion reflected the conclusions drawn from responses to the questionnaire, as described below.[4]

The Digital Picture's report also provides an overview of existing image database provision relevant to the arts education community,[5] which can be read in conjunction with WP2 and WP10-12 of the current document. In addition, the report summarises other relevant activities and reports:[6]

- The Democratising the Image round-table at the Computers and the History of Art (CHArt) Annual Conference (November 2005)
- The Compare and Contrast research project into digital images and art historians
- Creative Commons
- The Madison Digital Image Database (MDID)
- The Ministerial Network for Valorising Activities in Digitisation (MINERVA) and the Multilingual Inventory of Cultural Heritage in Europe (MICHAEL)
- The Museum Educational Site Licensing Project (MESL)

Information obtained from The Digital Picture's report has been supplemented by discussions with project staff. Whilst The Digital Picture is concerned with all aspects of the use of digital images within arts education, this summary focuses upon issues relating to image collections.

De facto, image collections played a large role in responses to The Digital Picture. This was accompanied by very real concerns at the impact which a growing use on digital images is having on research, educational processes and pedagogical values.

11.8.2 Demand for collections[7]

The community is currently frustrated by the difficulties of finding digital images they require: only 17% of respondents believe that existing digital image resources are sufficient for research purposes. Some of the reasons for this are outlined in the report: there are a great many existing collections, and

[4] § 3.1 on pp. 9-10, § 4.3 on p. 33.

[5] § 5.1 on pp. 40-48.

[6] § 5.2 on pp. 48-52.

[7] § 4.2.4 on pp. 18-19, § 4.2.6 on pp. 22-3, § 4.2.9 on pp. 29-30.

> It is patently unreasonable to expect everyone who wants an image, for whatever reason, to have to visit, subscribe and search each of these sites in turn until they happen to stumble on what they want. The opacity of just what is available is compounded by the different types of collections (e.g. JISC/AHRC funded, National Institutions, HE providers, private); the apparent need for owners to 'brand' their own collections; multiple, different authentication procedures; frequent lack of provenance; the range of costing models; poor or vastly differing search/navigation mechanisms; inconsistent metadata standards, when applied at all; protection of ownership rights; and the blurring of just what can be used, by whom and for what purpose.[8]

Existing image providers each offer part of the solution to the problem, yet these partial solutions often overlap: there is often a massive duplication of effort and resources. [9]

Whilst 91% think that finding images should be straightforward, only 37% feel the web offers the best solution, and supplementary responses suggest that this was due to both the narrow coverage and low quality of the images available there. However, some respondents felt that the serendipitous nature of web research could be stimulating.

A majority – 57% – of questionnaire respondents stated that they would like some form of web-based central repository – ideally provided via their institution (47%), and freely available. However, there was general scepticism that such a collection could secure the necessary cooperation and funding from the relevant stakeholders to ever be a working possibility. One respondent suggested that a register of the locations of relevant images might be a compromise.

Thus, although The Digital Picture report identified a surprising number of online databases of digital images, it recognised that the nature of this provision – fragmented provision, often of limited collections – is a major impediment to resource discovery.[10]

The notion of a central repository was also tempered by a general scepticism about any one collection's ability to either be aware of or provide all the images that any one individual might require. Consequently, 72% of respondents would also like to be able create their own images – although they remained aware of the intellectual property problems this might cause.[11]

11.8.3 Community involvement[12]

Any potential solution to the problems of current provision must be community led. According to The Digital Picture report,

> most people would now agree that, if a solution is to be found, it will need to be led by the needs of the users, no matter how complex such a resulting 'system' may seem. This view is already strongly recognised across the education sector, both within JISC and on a broader scale.[13]

This is one The Digital Picture's major conclusions, and its importance cannot be over-emphasised.

---

[8] § 2.1.2 on p. 5.

[9] § 2.1.2 on p. 5.

[10] § 5.1 on p. 40.

[11] See also § 2.1.2. on pp. 3-4.

[12] § 2.1.2 on pp. 5-8, § 4.2.10 on pp. 31-2.

[13] § 2.1.2 on p. 5.

Although the specific issue of community involvement was not raised in the questionnaire, several respondents – presumably reflecting the opinions of their colleagues – stressed how vital it was in their general responses to the consultation. One noted that they needed "Support for heterogeneous networks that allow teachers and researchers access to the hardware and software tools that are most suited to our needs, not those of the IT managers",[14] whilst another put the issue particularly clearly: "I'm pleased by your demand-driven, rather than supply-driven approach, since other initiatives don't seem to have learnt from the UKeUniversity mistakes."[15]

However, whilst the vocabulary of 'user needs' has been adopted by many online image providers, we have already seen how few respondents to The Digital Picture questionnaire feel that these providers actually meet their needs.

Thus, there is a problem with the ways in which community involvement is currently being incorporated into image provision. A significant aspect of this is the focussing of effort on specific, small communities at the expense of the broader arts education community – a problem exacerbated by the competitive nature of the current education culture.

But that culture is beginning to change: The Digital Picture notes the JISC Images Working Group (IWG)'s proposal to collate and make available those images needed by the community. The IWG's aims are paraphrased in The Digital Picture's report as follows:

- To provide the JISC education community with a long term digital image asset that is easy to use, free at the point of use, complies with common open standards, covers the broadest possible subject areas, is copyright cleared, is sustainable and supports the maximum manipulation by the user in support of fulfilling teaching and learning requirements.
- To create a fully operational managed service within the JISC Information Environment that supports a national image 'virtual reservoir'. The reservoir will provide the education community with the means to deposit and share their own images. Such a process will also enable the forging of alliances with non-education sector (e.g. galleries and museums) or commercial providers of image collections.
- Art and cultural heritage have a long-standing tradition of effectively using (and providing) image material for teaching, learning and research. The IWG believes the most immediate way to ensure the success of this proposal is through meeting the specific image requirements of these communities first, whilst exploiting the extensive images already available in parts of these communities. This might be done by funding institutions to digitise their copyright-cleared holdings, on the condition they provide copies of these images to the JISC National Art Digital Slide Library, and by forging alliances with such organisations as the Tate.[16]

In addition, such an initiative may include the national negotiation of access to subscription services, where these are deemed desirable by the community.

It should be noted that these proposals aim to enable or facilitate image creation: they are not focussed upon a 'portal', regardless of whether this may in fact be necessary. In addition, the IWG's proposal is not guaranteed success, which will only ensue

> if it can win the hearts and minds of a huge number of differing stakeholders; it will need to work extremely hard to be seen to meet the explicit needs of all

[14] § 4.2.10 on p. 31 (our emphasis).

[15] § 4.2.11 on p. 32.

[16] § 2.1.2 on pp. 6-7.

> the different users in its communities, and it will need to gain sufficient support from all the institutions, organisations and individuals who can supply the raw products, the images.[17]

Once again, community support emerges as the crucial factor.

The Digital Picture proposes that this might be secured by building a base corpus of commonly-used images, emphasising that it is only a starting-point for continued, community-led development. According to the report, this might incorporate a number of approaches:

- The JISC is already at the forefront of enabling digitisation of images for education – many JISC-funded collections could be made available to the reservoir
- Ongoing monies could be made available for the commissioning of specific photographs for known use in educational practice
- Continuing funding could be made available for the digitising of important collections
- Negotiations could take place with major National institutions to make their collections available through the reservoir, perhaps via OAI harvesting
- HE Institutions could be encouraged to provide access to local collections
- The reservoir could be capable of harvesting from institutional repositories
- Deals could be negotiated with commercial providers for nationwide access to images that are required
- Funding could be made available for the outright purchasing of specific images/collections
- Those who create images as part of their educational practice could be enabled to add their images to the reservoir
- Funding for digitisation could always come with a proviso that the results be made available (where feasible) to the reservoir – this notion could be encouraged within other funding bodies
- Digitisation bureaus could be established to scan or photograph image collections, either at specified places or as a 'mobile' service

The complex nature of the arts education community has been noted above. However, the precise basis on which communities are likely to be constituted has implications for the form of any community-based solutions, and the terms of which access may be granted to them. The reference, quoted above, to "the non-education sector (e.g. galleries and museums)" illustrates the problem. In addition to the many HEFCE-funded museums in HE institutions, the majority of museums consider themselves to be educational institutions – a view reinforced by the AHRC's decision to award 'academic analogue' status (and therefore eligibility for funding which originates with HEFCE) to museums based outside HE institutions.[18] Museums and their staff are potential contributors and users of images to community-based collections which reflect their particular interests; thus, they are very much members of those communities, regardless of whether their funding comes from 'educational' sources. Any subject-based community will necessarily include members who are not represented by organisations, such as the JISC, with remits restricted to Higher and Further

[17] § 2.1.2 on p. 7.

[18] http://www.ahrc.ac.uk/apply/research/mandg/academic_analogues.asp.

Education institutions. In the words of a representative of one such community, the Association of Art Historians, in its response to The Digital Picture,

> we also represent art historians who work in schools, museums, or as independent scholars. Digital images are crucial to the work of all our members, all of whom need unfettered access to high-quality digital images. We would urge The Digital Picture to press for mechanisms that ensure that all those involved in the teaching and research of art history, whatever their institutional status, have access to the digital resources they require.[19]

The Digital Picture also describes the diverse nature of the 'arts education community':

> As well as the obvious, simplistic, definitions, such as users from vastly different subject domains (graphics, fine art, art history, ceramics etc.) or those with fundamentally different roles (e.g. students, practitioners, lecturers or librarians), there are a multitude of other ways that users can differ from one another. For example, there are users from different areas of education (non-vocational, lifelong learning, bachelor's degree, post doctoral research etc.), or users with different learning needs (dyslexia, physical impairment, English as a second language etc.). Furthermore, we need to consider: how users of born digital images may differ from users of digitised images; distinctions between 'digital natives' – those who have grown up in the digital environment, and 'digital immigrants' – those for whom the digital environment has been a relatively recent introduction; and users of different types of images (e.g. bitmap and vector).[20]

In addition, individual image users can fall into different categories at different types, and their motives and actions are complex, inevitably shaped by external pressures such as institutional politics, inter-personal relationships, etc.[21]

Similarly, The Digital Picture report notes the international nature of any digital image collections that may be created for the arts education community: they will both call upon the resources of, and be of interest to, institutions and individuals in many countries, opening up the potential for international collaboration.

Given the widely-distributed nature of the relevant communities, and in the light of experience on The Digital Picture's 'virtual consortium', measures will need to be taken to ensure that those who are consulted in future truly represent the communities they purport to; and that all those who should be represented, are.[22]

One important point regarding the mechanics of community involvement emerged in the expert seminars. There was clear concern that centralised image repositories might threaten the livelihoods of slide librarians, whose expertise – the fruit of many years' experience – would be invaluable to the creation of digital image collections. It was proposed that they should become major contributors of cataloguing expertise and metadata quality-control.[23] In the words of the Association of Art Historians' response to The Digital Picture,

> Slide librarians (many of whom are AAH members) will play a crucial role in the process of transition, and must be supported with the training and budgets

---

[19] p. 55.

[20] § 2.1.2 on p. 4.

[21] § 2.1.2 on p. 5.

[22] § 4.4 on p. 40.

[23] § 4.3.3 on p. 39.

> required to perform their duties. Even after digitisation of analogue slide libraries, they will continue to perform a vital role in tracing or creating high-quality digital images required by students and lecturers, and in administering their storage and delivery, and the rights for their use. If major centralised repositories of quality digital images are created, slide librarians will play a crucial role in transmitting their users' needs to these repositories.[24]

11.8.4 Image quality[25]

Clearly, image quality is of great significance to the community: 82% stated that issues such as colour, contrast and clarity were important. However, notions of what constitutes sufficient quality were contested. In the words of The Digital Picture's report,

> heated discussions have been had about just what 'fit for purpose' means: to an art history lecturer, high resolution may be essential to make a particular point, whereas, for an FE student illustrating an essay, the subject matter may be more important than the quality of the image itself.[26]

Thus, the fact that currently 87% of respondents feel that they are excellent for talks and lectures, 74% consider them suitable for uses beyond web or intranet pages and 69% believe them to be useful for creating new artwork should not, perhaps, be given undue prominence, and several respondents expanded upon the ways in which digital images were currently generally not of sufficient quality or their needs.

Interestingly, one respondent suggested that the problem lies with current online collections: "The real deficiency is that too few museums and image sources provide images of high enough resolution, or detail."[27] Notwithstanding, 78% of respondents believe that an image must come from a reliable source.

The issue therefore seems to be one of fitness for purpose: the community requires images for a variety of end purposes, requiring different levels of quality.

11.8.5 Metadata[28]

Closely allied to the quality of digital images is the quality of the accompanying metadata. The community believes strongly that metadata is important: 62% feel that images must come with associated information. As one respondent put it, "Lack of metadata is what stops you from finding images."[29] Clearly, given concerns about image provenance noted above, this should include images' sources, allowing users to verify their reliability for themselves. The importance of metadata for students to cite images accurately was stressed, as was its importance in identifying rights holders and any restrictions.

11.8.6 Intellectual property issues[30]

Intellectual property was (just) the subject of greatest concern to respondents. The community covered by The Digital Picture includes both creators and users of images, and responses to the survey embody the concerns of both groups: 68% acknowledge the need to protect the financial rights of image creators, whilst 75% believe the use of images should be

---

[24] p. 58.

[25] § 4.2.2 on pp. 14-15, § 4.2.6 on pp. 22-3, § 4.3.1 on p. 34.

[26] § 4.2.2 on p. 14.

[27] § 4.2.2 on p. 15.

[28] § 4.2.6 on pp. 22-3, § 4.3.1 on p. 34.

[29] § 4.2.6 on p. 23.

[30] § 4.2.7 on pp. 24-6, § 4.2.10 on pp. 31-2, § 4.3.1 on p. 34, § 4.3.3 on p. 39.

free within education. There seems, therefore, to be little conflict between the two groups, with creators largely being willing to allow their work to be used for free for educational purposes – duly acknowledged.

The issue, however, is what constitutes 'educational use': for example, one respondent noted that they would be happy for their images to be used in a lecture or school project without charge, but would expect remuneration for their use in a higher education course pack. This might be addressed by creating a generic licence (along the lines of the Creative Commons scheme[31]) defining and allowing 'educational use'.

Several respondents felt that the issue was one for government, noting that it has the power to exempt education from copyright restrictions if it wishes to do so, and that institutions such as galleries and museums should be included amongst those institutions currently defined as 'educational establishments' under current UK legislation (presently restricted to schools, colleges and universities).

The vast majority of respondents did not wish to be bothered by legal issues themselves: 79% believe their institution should take care of them. However, a notable minority felt that individuals should be aware of the issues, either because institutions could not police their staff effectively, or because they felt students needed to be educated about the importance of copyright. Provision of clear rights metadata with each image (see above) would address both these concerns.

In addition, copyright was a significant concern at the expert seminars, where delegates recognised that existing copyright legislation in this area still needs to be tested in court. Any rights clearance strategies would have to be assessed in terms of their potential risks until this happened. Many delegates felt that nothing short of a change in the law to allow free dissemination of copyright images for educational purposes would work as a solution to these problems.

11.8.7 Funding[32]

The Digital Picture identified lack of funding for digitisation as a significant issue: 71% of respondents believe that more money needs to be spent.[33] When money was available, it was usually dedicated to essential work and equipment, or to innovative projects, not to the straightforward task of transferring existing analogue resources into digital formats.

11.8.8 Potential solutions[34]

In order to focus discussions at the expert seminars, The Digital Picture proposed a series of potential solutions to the issues raised in the questionnaire, centring around different models for the provision of digital images. These were then discussed at the seminars.

National model – external

A single, centralised collection made available to those requiring access. This could incorporate an image viewing and presentation tool.

"There was fear that, in adopting this model, control and power would most likely be appropriated by outside organisations, who had different, or even opposing motives when undertaking their core work", as well as concern that lack of consultation over the creation of the corpus of works to be included would lead to the digitisation of irrelevant material. "While

[31] http://creativecommons.org/

[32] § 4.3.1 on p. 33.

[33] § 4.2.8 on p. 27.

[34] § 4.3.2 on pp. 35-8, § 4.3.3 on pp. 38-9.

the benefits of centralised and managed control were appreciated, the fear of losing control and resources to an 'external body' was strongly voiced."

National model – internal

A federated collection based upon harvesting and/or interoperability of discrete, local collections, with a single point of contact for searching and downloading, and uploading or exposing, images. This could incorporate an image viewing and presentation tool.

Very much the preferred option:

> The idea of joining up the resources that already exist in [the] community was deemed to be an excellent and the most preferred starting point for many delegates. ... Perceived ownership by the community, and a process with which the community could engage and interact, was felt to be paramount, and essential for community building and the success of such a strategy.

However, such a model would still require a centralised agency to set standards for quality, metadata, rights clearance, etc., and to monitor compliance with those standards. It was proposed that the model could begin work with a comprehensive audit of the resources and owners of existing digital and analogue images. In addition, "There was a general consensus that for this model to work, and to attain the correct kudos, uptake and collaboration should be secured from all the National museums, whose trove of public domain works of art should be made accessible through such a system."

Local model

Discrete, local collections housed and disseminated within particular institutions, and perhaps made public via individual websites.

Whilst the local model "was felt not to offer the breadth of opportunity that the national model offered", practical and institutional considerations "may make this solution favourable for some". Its real potential, for many delegates, lay in the ability of local collections to interact with a national model. In addition, if individual institutions approached the creation of collections independently, "this would breed a lack of consistency, duplication of effort, and an inefficient use of limited digitisation funds".

Commercial image libraries

Subscription-based access to existing image libraries (n.b. this option includes resources which, although not-for-profit, adopt a subscription/charging model).

> While the benefits offered by 'commercial' image libraries ... were universally recognised at the seminars, it was felt that such 'core' resources could only be usefully used to augment a more internal model; it was thought that the ability to select works from such libraries would be better and more cost effective than subsidising the use of the whole collections, in order that the community could still lead on the selection of digital content, without being dictated to by outside agents.

In addition, any time restrictions placed on the images supplied by such libraries were felt to be unworkable. Whilst the tools and functions provided by such collections were desirable, any lack of interoperability of systems or images would greatly restrict the usefulness of such collections.

National fund

Not a collection model, but a proposal to establish a fund dedicated to the creation of digital images for the arts education community at a range of levels.

> It was acknowledged that there was a need to attract more money for the digitisation of still images, and that other object types, like GIS and Virtual Reality, seemed to enjoy more success with securing funds. ... [C]ontinuity, with a long-term, continuous funding stream was felt to be the only way to address the issue, which would be ongoing for the foreseeable future. Some kind of national dedicated fund for digitisation, with published strategic priorities, would aid all institutions to position themselves appropriately, and assess their ability to contribute, without wasting time and resources on building unsuccessful bids.

Serendipity

Using existing, freely-available resources such as Google's image search.

Google and other search engines remain highly successful, and any solution must emulate such systems' apparent functionality and ease of use. However, the "lack of metadata, provenance, proper rights clearance, integrity or quality" inherent in such systems were unacceptable in an academic context. (Although the recent introduction of a Creative Commons filter on Google advanced searches will go some way to addressing one of these issues.[35])

Other solutions

Additional solutions proposed during the expert seminars included peer-to-peer and file-sharing solutions using tools such as Napster or Flickr as models.

There were also several suggestions for useful features to be incorporated into any adopted model:

- An Amazon-like system for automatically recommending other resources and images, based on one's search results
- Use of wikis for annotation of images by users (subject to meeting concerns about quality-control and integrity)
- A grading system for image quality, to be published alongside the image
- Assignment of concepts, themes and emotions as indexing terms, alongside more conventional subject-headings
- Methods for preserving users' searches and routes through the system for future re-use
- The ability to upload new content

Summary

There is a clear preference for the 'National model – internal'. By its very nature, this will be built upon devolved collections which use the 'Local model'. Any solution will require the creation of a 'National fund' to meet the additional costs incurred in digitisation.

11.8.9 The Digital Picture's conclusions

There is a clear demand in the arts education community for a greater commitment to provision of digital images. However, such commitment must respond to the community's needs. There is a demand for community-led image repositories, and in order to meet the community's needs they must:

---

[35] E.g. http://www.google.co.uk/advanced_search?hl=en; see http://www.google.com/support/bin/answer.py?answer=29508 for further details.

1. provide one-stop access to federated resources
2. incorporate locally-produced collections of images
3. allow institutions to house their own digital image collections
4. allow institutions to share image collections
5. allow users to deposit images
6. allow other public sector image collection owners to add their collections
7. have an acquisitions policy shaped by the community
8. create a critical mass of images relevant to the community's needs
9. contain a broad spread of material, avoiding the fine arts focus of most current repositories
10. incorporate existing public-domain images
11. aim wherever possible to acquire images of the highest possible quality, repurposing them as required
12. provide users with continuous online access
13. allow for serendipity
14. identify the level of quality and uses for which each image is suitable
15. ensure that images are provided with an agreed level of metadata, which must cover the images' identification, their provenance, and their copyright status
16. define acceptable 'educational use'
17. try and acquire images with unrestricted rights for 'educational use'
18. state clearly and simply the copyright status and permitted uses for all images
19. secure sufficient funding and commitment from stakeholders

This would best be implemented as a federated collection based upon harvesting and/or interoperability of discrete local collections, and would need the support of dedicated, long-term funding.

The Digital Picture has noted the JISC Image Working Group's proposals for community involvement and investment in facilitating and enabling future image provision, and has also proposed creation of a corpus of frequently-used images to help secure community support for these initiatives. In order to meet existing community concerns regarding quality, metadata, and intellectual property rights, it also proposes:

20. raising awareness of good digital image practice (e.g. quality, metadata standards, copyright)
21. encouraging image creators to make use of extended ownership mechanisms (e.g. Creative Commons)
22. ongoing negotiations with CLA and DACS to increase the potential provision of scanned images[36]

[36] § 2.1.1 on p. 8.

11.9 A Survey of the provision of digital images by museums

Rupert Shepherd

11 January 2006

11.9.1 Introduction

This report is based on an examination of the museum-based online image collections listed in, together with an examination of other leading museums and art galleries in the U.K. (particularly nationally-funded collections) and a few significant museums abroad. A full list, with URLs, is provided at the end of this report. Representatives of the following museums and museum projects were interviewed in greater detail between 1 and 14 December 2005 in order to establish the reasoning behind current provision and attitudes towards providing material for possible future community-led collections:

Bolton & Bury Treasures in Trust

COMPASS (The British Museum)

Ingenious (The National Museum of Science and Industry)

The National Galleries of Scotland

The National Gallery

The National Portrait Gallery

Tate

11.9.2 Availability of image collections

There are a great many museum-based websites which present images online. In large part, this is due to continued government pressure to make collections accessible to the broadest possible public, both physically and intellectually, as embodied in projects such as NOF-Digitise, which

> ... provided funding to make learning materials available, free of charge, on the Internet. From programme inception in April 2000, it ran for 48 months until March 2004 when the final digitisation projects were signed off. The programme has unlocked the learning resources of libraries, archives, museums, galleries, colleges and universities, charities, voluntary organisations and others by converting them into electronic form. The total allocated budget for the programme, £50 million, was spread over 150 discrete projects consisting of various partnerships between a total of nearly 500 diverse organisations throughout the United Kingdom.37

However, this does not mean that presence of an image collection on a museum website can be taken for granted. The National Museums of Scotland, for example, have no images on their website, but have contributed to SCRAN.38 Other museums appear to have made no provision at all to display their images online – including national collections such as the Wallace Collection and Sir John Soane's Museum.39 Of the great European art collections, the Uffizi, the Staatliche Museen zu Berlin and the Prado do not systematically provide online

[37] http://www.ukoln.ac.uk/nof/support/#briefNOFhistory. A searchable portal to the collections created under NOF-Digitise is provided under the name of Enrich-UK at http://www.enrichuk.net.

[38] http://www.nms.ac.uk/nms/home; http://www.scran.ac.uk

[39] http://www.wallacecollection.org; http://www.soane.org

images, whilst the Kunsthistorisches Museum in Vienna only provides small images of a small selection of its works.40

11.9.3 Discovery of image collections

Image collections are not always easy to discover. As the NOF initiative suggests, funding tends to be granted to discrete, small-scale projects, rather than large-scale, cross-institutional ones.41 When projects are organised across institutions, they are usually tightly focussed on specific geographic regions or subjects: very few (notably SCRAN) provide comprehensive coverage at a national level.42 Assuming one is looking for an image of a specific object, the logical space to start would be the website of its owner. However, many image collections are placed under different titles and on different URLs. Image collections may be divided up into different projects or collections within the institution (the Ashmolean Museum, which presents twelve different image-rich resources online, with no cross-searching, is a notable example), and may be buried deep within the collection's website in different areas.43 Image collections are often not regarded as such, but as aspects of more general online resources, a point which representatives of individual collections were anxious to emphasise. Consequently, it is not always clear whether the links one follows will in fact lead to collections of images. Once they have been found, a surprising number of image collections do not have search facilities, but offer the opportunity to browse using a variety of categories – a facility of varying usefulness, if choosing to browse the 20th-century objects might produce a list of more than a thousand images.

11.9.4 Content of image collections

Selection

As noted above, very few institutions provide images of their entire collection online: the National Gallery, National Portrait Gallery and Tate are the major exceptions.44 Significantly,

---

[40] http://www.polomuseale.firenze.it/uffizi; http://www.museen-berlin.de; museoprado.mcu.es; http://www.khm.at

[41] Fragmented funding is a significant complaint of those involved in creating digital resources: *Digitisation in the UK*, p. 4.

[42] http://www.scran.ac.uk. As *Digitisation in the UK* notes in § 1.1 (on p. 4), 'There is no UK register to map individual digitisation projects and therefore no authoritative resource to aid discovery and prevent duplication', although resources such as Enrich-UK (http://www.enrichuk.net) provide registers of, for example, NOF-funded projects.

[43] The Ashmolean's homepage is at http://www.ashmolean.museum; its various image collections and their URLs are: *The Burgon Archive*, http://www.ashmol.ox.ac.uk/ash/amulets/burgonarchive/; *Oxfordshire's Historic Archives*, http://www.ashmol.ox.ac.uk/ash/amps/oha/; *Iraq: Navel of the World*, http://www.ashmol.ox.ac.uk/ash/amps/iraq-navel/; *Ancient Near Eastern Terracottas*, http://www.ashmol.ox.ac.uk/ash/amocats/ANET/; *French and Russian Drawings*, http://www.ashmolean.museum and follow link to AMOS; *The Paintings Catalogue*, http://www.ashmol.ox.ac.uk/php/am-search.php?db=wapaintings; *The Creswell Archive*, http://creswell.ashmol.ox.ac.uk; objects from the Department of Antiquities' collection of brass rubbings, http://www.ashmol.ox.ac.uk/ash/departments/antiquities/brass/; *John Evans' Lake-Dwelling Collection*, http://www.ashmol.ox.ac.uk/ash/amps/jevans/; *The Elements of Drawing*, http://ruskin.oucs.ox.ac.uk; *PotWeb*, http://potweb.ashmol.ox.ac.uk; *TileWeb*, http://tileweb.ashmolean.museum (which also includes objects in Worcester College).

44 http://www.nationalgallery.org.uk/collection/default_online.htm; http://www.tate.org.uk/servlet/BrowseGroup?cgroupid=999999956. The Louvre's *Atlas* database includes all objects which are on public display: cartelfr.louvre.fr/cartelfr/visite?srv=crt_frm_rs&langue=fr&initCritere=true. The immense mass of undigitised material is noted in *Digitisation in the UK*, p. 2, which also states that 'Significant gaps in provision remain in many disciplines, including those seemingly well served' (§ 1.1 on p. 4).

the former has collections totalling about 2,500 objects, the latter two, less than 100,000 objects each. As collections can easily run into hundreds of thousands of objects or more, this is hardly surprising, but it does raise the question of selection.

Criteria for which objects have been placed online are seldom explicitly stated. Conversations with individual institutions have revealed that selection decisions are usually made by curatorial staff, based upon

- whether images already exist
- whether works are on display
- the perceived popularity of images in the collection (the "greatest hits")

The V&A, which does describe its selection policy, states that its online images represent a combination of recently-published works, and works which are less-well-known or not normally displayed.45 In the majority of cases, there is little evidence of audience consultations or the use of methodical surveys in order to determine which images should be made available.46

Image size

The majority of images are of comparatively small size – few collections present the majority of their images at more than 600-700 pixels on their longest sides. (To adopt the vocabulary established in WP10, collections are limiting deliverables.) Larger images are usually presented using 'zooming' technologies which only present a section of the entire image.47 The National Gallery goes so far as to combines zooming technology with visible watermarks within the zoomed details.48 SCRAN, which includes the holdings of many Scottish museums, only makes thumbnail images available without a subscription. There are a few notable exceptions: the Eadweard Muybridge Bequest, digitised by the Kingston Museum and Heritage Service, and Epact: Scientific Instruments of Medieval and Renaissance Europe, a collaborative project run by the Museum of the History of Science at Oxford, include images over 1200 pixels on their longest sides;49 and, most impressively, the Fine Art Museums of San Francisco present tiled images (which therefore cannot easily be downloaded) which assemble to create images more than 2600 pixels high.50 The Rijksmuseum in Amsterdam includes a substantial collection of images at 1600 pixels.51

---

[45] images.vam.ac.uk/ixbin/hixclient.exe?_IXSESSION_=&submit-button=search&search-form=main/index.html

[46] A shortcoming addressed by recommendation 1.2.3 of *Digitisation in the UK* (see p. 5); see also § 3.4 on pp. 13-14.

[47] E.g. *Revolutionary Players*, http://www.revolutionaryplayers.org.uk/, *Exploring the Potteries*, http://www.exploringthepotteries.org.uk/, and *IMAGINE*, http://www.imagine.org.uk/, which zooms images to a remarkable level of detail.

[48] Accessible via searches conducted at http://www.nationalgallery.org.uk/collection/default_online.htm.

[49] 213.48.46.171/museum/muybridge/ and http://www.mhs.ox.ac.uk/epact/. Collections which produce images between 700 and 800 pixels in length include the Courtauld Institute's *Art and Architecture* site, http://www.artandarchitecture.org.uk; the V&A, http://www.vam.ac.uk/collections/; and the Louvre's *Atlas*, cartelfr.louvre.fr/cartelfr/visite?srv=crt_frm_rs& langue=fr&initCritere=true.

[50] http://www.thinker.org/fam/about/imagebase/index.asp

[51] These are from the Museum's online 'Educational Collection' of 10,000 objects: http://www.rijksmuseum.nl/collectie/zoeken/. In addition, they provide 580-pixel images from their Collections Management System records of 50,000 images, although that collection is only published in Dutch.

The size of the majority of images (≤ 600 pixels on the longest side) raises clear problems for their reuse. A horizontal image of this size will fill the screen on a VGA monitor at 640 x 480 pixels, and a vertical one an SVGA monitor at 800 x 600 pixels. However, monitor and data projector resolutions are increasing: these images already seem small on SVGA monitors set to 1024 x 768 pixels, and, as XGA monitors at 1280 x 1024 and 1400 x 1050 become more widespread, will seem increasingly unsatisfactory. This is even more true of data projectors which, although lagging behind monitors in pixel dimensions, are the preferred means of delivery for images within the classroom. A clear 600-pixel image at today's screen sizes may still be acceptable as a general illustration, but it simply does not contain sufficient information to be the focus of any kind of detailed attention and investigation. Again, it will serve as an adequate small printed illustration, but not if any significant information is to be gleaned from it.

Metadata

The level of metadata provided by museums is variable. It almost invariably focuses upon the image content, not the image itself. It tends to fall into two types: a summary catalogue entry, or framing text. Catalogue entries give the basic information about an object (usually based on a core of object type, title, maker, date, materials/techniques, location). Framing text may describe or interpret the image's content, or may provide a context which the image is used to illustrate. It tends to be aimed at the general reader, and so its usefulness depends upon the purposes for which the image is required.52 Use of embedded metadata (notably IPTC/XMP) does not seem to be widespread – presumably in part because it would increase the relatively small file-sizes, and in part because the files are not intended for circulation or reuse, and so there is no perceived need for metadata to travel with the object.53 Dublin Core metadata may be present, but is not explicitly labelled as such. Image capture and processing metadata, whether EXIF for direct photography or any other form, seem practically non-existent.54

Images are usually served dynamically, making it difficult to cite or refer to them directly: URLs are either too long to be manageable, or unstable.55

Enhancements

Collections provide various enhancements in order to increase the attractiveness (and therefore the 'stickiness') of their websites. Beyond the additional content on the surrounding websites, some collections offer participation in the online community of which they are part.56 Many sites offer the facility to create folders of objects, and several to view zoomed details of works in the collection. Print-on-demand, where hard copies of images are ordered and paid for online, printed and despatched to users, are a very popular enhancement with the wealthier institutions.57 Other are more specific: Turning the Page, which reproduces the

---

[52] The importance and cost of metadata creation is noted in *Digitisation in the UK*, p. 4.

[53] Significantly, where embedded metadata is used – the National Museum of Science and Technology's *Inenious*, http://www.ingenious.org.uk/See/, and the Smithsonian Institution's *Smithsonian Images*, smithsonianimages.si.edu/siphoto/siphoto.portal – the resources are in part based upon commercial picture.

[54] Some more recent Smithsonian images carry EXIF metadata.

[55] Two exceptions: MOMA, New York, provides an option for users to create a stable URL for the image they are viewing (http://www.moma.org/collection/search.php); and the National Portrait Gallery provides stable URLs for each entry in its online catalogue (http://http://www.npg.org.uk/live/collect.asp). Those collections which allow users to create and publish folders of images provide a more cumbersome means of referring to online images.

[56] E.g. Tate, http://www.tate.org.uk/onlineevents/, and *Ingenious*, http://www.ingenious.org.uk/Debates/.

[57] Offered by *COMPASS*, the National Gallery, the National Portrait Gallery, *Art and Architecture*, and *Collage*.

effect of leafing through a bound volume;58 glossaries; random image generators; 59 and content-based image retrieval.60

11.9.5 Audiences and intended uses

Although many online image collections reflect institutions' educational roles (and the priorities of funding schemes such as NOF) by providing educational resources, the great majority do not seem to be intended for further or higher education audiences. The images are too small to be the subject of serious examination, provision of metadata is variable and separated from the image file, they cannot easily be cited and it is not always easy to locate a specific image. This observation was borne out by interviews with museum representatives, all of whom stated that they were addressing a general audience, albeit usually a comparatively well-educated one.

Consequently, routes into the collections tend to include processes to narrow down broad concepts into a manageable list of objects to browse, or methods aimed at non-expert users such as map-based interfaces. 61 The data that is provided is basically intended for on-screen use, and not for saving and retrieving later (although some collections, such as the Courtauld's Art and Architecture, do provide systems for saving collections of images for later re-use62). In the majority of cases, the assumption seems to have been that users will find an image, view it onscreen, and move on – that the data is effectively 'disposable' as far as the user is concerned. Again, this is borne out by the copyright statements on the sites, which focus on online viewing of the images. Although there have been few, if any, formal surveys of what visitors to museum websites have actually been doing with images, anecdotal evidence suggests that users have also intended to download them for class-room use, print them out for inclusion in project work, and on occasion reuse them in their own websites.

Where collections are presented as a large mass of searchable images, this is sometimes because the viewer is in fact being presented with a website from the organisation's picture library – in other words, a collection of images primarily aimed at picture researchers working for publishers, magazines and newspapers, packaging designers, advertisers, etc.63

11.9.6 Rights

In addition to limiting deliverables, collections usually also adopt the educational approach to rights management defined in WP10, with statements – usually on a separate web-page, linked to from within the collection – outlining what may or may not be done with images. The conditions attached to use of UK-based collections are fairly standard: viewing, downloading and printing for personal use are acceptable, but commercial use or publication are forbidden. Educational use is sometimes permitted, although seldom defined. The Victoria & Albert Museum is unusual in inviting 'creative professionals' to use images for paste-up

---

[58] Offered by the British Library: http://www.bl.uk/onlinegallery/ttp/ttpbooks.html.

[59] Both at Tate: http://www.tate.org.uk/collections/glossary/; http://www.tate.org.uk/collection/carousel/

[60] Offered by the Hermitage: http://www.hermitagemuseum.org/fcgi-bin/db2www/qbicSearch.mac/qbic?selLang=English

[61] Solutions adopted by, for example, *Pastscape*, provided by English Heritage's National Monuments Record, http://www.pastscape.org/homepage/; and *Windows on Warwickshire*, the result of a partnership between the Shakespeare Birthplace Trust, Warwick Castle, Compton Verney, Nuneaton & Bedworth Museum & Art Gallery, Warwickshire Museum, Warwickshire County Record Office, and Warwickshire Library & Information Service – see http://www.windowsonwarwickshire.org.uk.

[62] http://www.artandarchitecture.org.uk; several other collections offer this facility.

[63] As is the case with the National Trust: http://www.ntpl.org.uk, as well as *Smithsonian Images*, cited above.

work.64 The majority of US-based collections are somewhat more generous, as they have to incorporate the 'fair dealing' provisions of US copyright law: use for school websites, criticism, commentary and news reporting is often explicitly allowed.65 However, commercial use – defined as use in any resource that charges for services, or that takes advertising – is explicitly excluded.

11.9.7 Summary of provision

To summarise, image provision from the museum sector seems to be top-down: the institution decides who its audience is and how it can best serve their needs. Consultation with communities of users regarding the selection of material to reproduce, image sizes and formats, and the data which accompanies the images has not been extensive. The target audiences are usually an educated general public. The end result is disposable data.

These conclusions raise three questions:

1. Which communities might use museum websites as sources of images?
2. What are their requirements?
3. Are museums able or willing to meet these demands?

11.9.8 Museum image collections and CLICs

Interested communities

For the purposes of the current report, answers to question (1) will be restricted to further and higher education. As museums house objects relevant to a great many disciplines, there seems little point in trying to narrow down the subject communities for which their holdings are relevant; individual communities will be best placed to determine which museums' holdings are most relevant to their interests.

Interested communities will include the great majority of those involved in these sectors: undergraduate and postgraduate students, researchers and lecturers, as well as support staff, notably librarians (including visual resources librarians).

However, it should be noted that many subject-based communities expand well beyond further and higher education, to include independent scholars, part-time and casual lecturers, school-children and teachers, interested amateurs, and even many museum staff.66 Community involvement in museum image collections will have to include these constituencies as well as those based solely in further and higher education.67

Sector requirements

Communities' requirements are related to the notion of fitness for purpose. Unfortunately, the purposes for which images may be used across further and higher education are too numerous to examine here and are, again, best left to individual communities to decide. Different subject-communities will have very different requirements. But it can be noted here that, although 600 pixel images of a generic subject will be more than adequate as aides-mémoire or reference illustrations in a student essay, other users will require images of specific (and often obscure) objects distributed at much higher standards than are currently provided by the vast majority of museums.

---

[64] http://images.vam.ac.uk/ixbin/hixclient.exe?submit-button=search&search-form=main/terms.html

[65] The National Gallery of Art (Washington DC) is the most stringent: http://www.nga.gov/copyright/index.shtm.

[66] A point made in the Association of Art Historians' response to *The Digital Picture*, p. 55.

[67] Although it is a subscription service, SCRAN, http://www.scran.ac.uk, offers rates for individuals, museums and schools – though not for freelance academics, unless they wish to pay a full commercial rate.

Meeting sector requirements

Broadly speaking, there is a managerial commitment amongst many public-sector organisations to make information widely available. In the words of the recent report on the potential use of Creative Commons licences by members of the Common Information Environment ('the CIE-CC report'),

> CIE and other public sector organisations wish to ensure that the resources they produce are used and reused as widely as possible. Many of these resources are useful for education, from primary schools to universities, for the public, through local history and similar interest groups, and will be of use to museum and library staff. By making resources available, and encouraging both reuse and redistribution, public sector organisations can encourage innovative and exploratory use of their material, much of which would not take place otherwise.68

Communication between museums and educational/academic communities

Museum-based image collections are currently directed at general users, although with a presumption that educational use will be made of them. However, there is little evidence of direct consultation between museums and Higher and Further Education institutions regarding those institutions' digital image requirements. Similarly, there is little evidence of direct consultation between museums and specific academic communities regarding the communities' image requirements. Thus, there is no existing mechanism for coordinating the digitisation activities of museums with the technical, metadata and content requirements of Further Education, Higher Education and academic communities.

Addressing the communication gap

Clearly, there is a need for some mechanism to fulfil this role. This will need to be able to accommodate the different requirements of different subject-based communities, as well as the many hundreds of museums which have the potential to contribute digital images.

These multiple interests are probably best served by the creation of two bodies:

1. A community/museum liaison committee. This would contain representatives of all communities which felt that museums could provide images relevant to their needs, and representatives of the museum community. On the museum side, this would best be run through one of the national organisations: the Museums, Libraries and Archives Council (MLA) and equivalent organisations in the devolved administrations, and/or the Museums Association. The Museums Copyright Group should probably also be represented. This would draw up and prioritise community requirements and pass them to:
2. A museum digitisation committee, which would establish the best methods of meeting community requirements. The number of museums which might be involved suggests that a two-tier structure might work best: a national committee, comprising the museum representatives from the community/museum liaison committee and a member from each of the UK regional museum hubs;69 and a series of local committees, one per region. These would take community requirements from the community/museum liaison committee, determine which museums should contribute content and how they would do so, and pass the relevant requirements to those museums.

Both bodies would have a role in allocating any funding in accordance with community requirements and museum needs.

This framework would also provide a solid structure for assembling subject-based core corpuses of digital images.

---

[68] Barker et al., p. 10.

[69] 'Hubs' were established under the MLA's *Renaissance in the Regions* programme: see http://www.mla.gov.uk/action/regional/00renaiss.asp.

Institutional concerns

However, many institutions are loath to make their collections freely-available, for various reasons. When asked if they would be willing to make their image collections available to some form of aggregated collection for educational use (either a unified collection or a virtual collection compiled by cross-searching exposed metadata), the majority of museum representatives responded favourably in principle, but noted that many issues would have to be addressed in practice before any commitment could be made. Their main areas of concern were:

- Loss of contextual information, and the wish to present their collections as a unified site rather than a series of fragmented resources
- Loss of intellectual control over their images and accompanying data
- Loss of physical control over images and their subsequent reuse
- Inability to grant rights to reproduce all their collections, due to works still being in copyright
- The potential for conflict with the museums' commercial activities
- The time and money required to prepare images and data for deposit or exposure
- Local infrastructure problems

This is confirmed by the CIE-CC report, which states:

> ... public sector organisations will have to make more outputs available to be used in more ways by people outside their organisations.
>
> However, public sector organisations are also under pressure to generate commercial revenue from their activities. This means that they would not want to publish resources in such a way that could damage their perceived market, so appropriate procedures for deciding what should be made available and under what conditions are necessary. The situation is also complicated by a requirement for public sector organisations to act in a non-discriminatory manner.
>
> In addition, resources produced by the public sector may have parts which are owned by third parties, or they may act as custodians of material which is entirely owned by third parties. In the event of mistaken publication of these resources to a wider user group than the third party agreed to, it is clear that there would be a greater risk of potentially embarrassing and/or costly disputes or court cases. There are also situations in which third party material has been licenced to the public sector body for use in a particular sector, for a particular geographic region, or for a specified period of time. Accordingly, selection procedures for deciding which resources can be made available and rights clearance methodologies would have to be further developed for resources which already exist and contain third party material. On the other hand, for future commissions of third party material there would be a positive advantage from having a standard set of rights to be assigned by the third party.
>
> There are also situations where identifying the rights holder of third party material is difficult, so-called "orphan works". ...
>
> There were also some concerns about potential damage to rights holders' reputation which could result from derogatory use of resources or activities which endorse an activity which the rights holder objects to. Associated with this concern about reputation were comments about ensuring that the most up-to-date versions of resources are used and that quality is maintained when resources are moved.

> Most rights holders accept that users should be able to have all rendering rights (print and display) and that they should be allowed to carry out certain acts of modification such as aggregating and embedding the entire resource with other resources. Other modification actions such as annotating, extracting and making actual changes to pictures and texts received more divided answers. While some believed that such modifications were essential others did not. Views about modification depend, to some extent, on the type of resource. In particular images were felt to be a resource type which ought to be protected from modification. An element of choice enabling rights holders to make individual decisions about whether each resource should or should not be modifiable is indicated.
>
> With regard to constraints on use it is clear that rights holders would expect to be attributed for producing resources and that they would like to be contacted if a user intends to make commercial gain from the resource. There were also a significant minority who considered technical protection and limiting the duration of use to be important. Restrictions based on geography and number of users were not considered to be so important, although there were some notable exceptions where existing resources are already licenced with geographic or sector restrictions.
>
> Transport rights must be granted by rights holders if they are to be reused. From the workshops, the attitudes of rights holders seemed to be divided as many were not comfortable with users giving or lending the resources. However it was noted from the workshop discussions that many rights holders would be prepared to relax restrictions on use if certain conditions of use are adhered to. Further evidence of this is referred to in section 4.5 [of the CIE-CC report] which explains how many of the rights holders were prepared to allow redistribution provided that certain conditions such as attribution and a restriction to non-commercial use are adhered to. 70

To summarise, image owners are concerned about:

1. Losing the intellectual context for their images and data
2. Maintaining the most up-to-date version of digital objects, and potential degradations in quality as they are reused
3. The possibility of the unrestricted recirculation of digital objects
4. Allowing digital objects (particularly images) to be altered (certain owners only)
5. The difficulty of clearing third-party copyrights in parts of their collections to allow for further dissemination, and of identifying rights holders for ‘orphan works’
6. Potential loss of opportunities to generate revenue from their collections
7. The resources required to make collections available
8. The danger of derogatory use of collections, or their use implying and unsuitable endorsement
9. Desire for attribution
10. Retaining the ability to situate collections behind technical protection, and/or to time-limit their use (certain users only)

Addressing institutional concerns

[70] Barker et al., pp. 20-21; see also Appendices, pp. 7-10.

The question remains: can these concerns be addressed easily? The adoption of three basic solutions would seem to go a substantial way towards doing this:

1. Rather than create a single repository for digital images, aggregate data from individual image owners' collections.
2. The conclusions of the CIE-CC report are that, in many (but not all) cases, the application of Creative Commons licences and one or two newly-created but standardised licences would address many of these concerns.
3. Funding will be required
    a. to fund the exposure of existing collections. This would, however, be a one-off, as the mechanisms and metadata required for exposure could be incorporated into workflows for the creation of future collections.
    b. to assist with rights clearance in new collections, where this is a major obstacle to digitisation.

To take the issues listed above one-by-one:

1. Losing the intellectual context for their images and data

An aggregated collection could ensure that users were led back to the owners' original site at a certain point in the search/delivery process.

2. Maintaining the most up-to-date version of digital objects, and potential degradations in quality as they are reused

Again, an aggregated collection would ensure that updates could be made easily by the objects' owners.

3. The possibility of the unrestricted recirculation of digital objects

It appears that, for many rights holders, this is not a significant issue if objects can be recirculated under agreed conditions (as defined by a Creative Commons or other licence).71

4. Allowing digital objects (particularly images) to be altered (certain owners only)

Alteration can be specifically prohibited using Creative Commons BY-ND or BY-ND-NC licence (attribution, no derivative works; or attribution, no derivative works, non-commercial); otherwise, it is allowed by all other Creative Commons licences.

5. The difficulty of clearing third-party copyrights in parts of their collections to allow for further dissemination and of identifying rights holders for 'orphan works'

Rights clearance will require administration and research, and will require funding to carry out for existing collections. Identifying rights holders will require funding. It should, however, be part of the workflow for any digitisation project, and so the decision to digitise should already have allowed for this expenditure in all but the most complex cases.

6. Potential loss of opportunities to generate revenue from their collections

The issue of commercial use is perhaps the most important. Many owners are happy to circulate their collections for educational use.72 But this is problematic: different owners are likely to define educational use in different ways or (perhaps because of funding) restrict use access to certain educational sectors. Such restrictions cannot

---

[71] Barker et al, p. 21.

[72] Barker et al., Appendices, p. 73.

be addressed by generic licences.73 However, it seems that the main reason for owners to wish to restrict access to educational users is to avoid commercial exploitation of their collections. In other words, a licence which prevents such commercial exploitation, such as a Creative Commons BY-NC licence, would seem to address some of these concerns. However, even with a no-commercial-use Creative Commons licence, it remains to be seen whether institutions will be willing to place reproduction-quality (i.e. commercially-exploitable) images on relatively open access. The CIE-CC report suggests that this is unlikely,74 presumably due to concerns over piracy in contravention of the licence. However, placing an educational collection behind some sort of technical protection (see 10 below) might alleviate these concerns. In addition, the non-commercial nature of educational institutions is by no means certain in today's age of tuition fees and charging for online continuing education courses.

7. The resources required to make collections available

As mentioned above, funding will be required to expose existing collections, primarily to ensure that images and metadata are formatted to required standards. (The necessity of standardising metadata is a significant question, addressed elsewhere in the CLIC report.) However, much of the work, once metadata crosswalks and so on have been established, can be incorporated into workflows for future digitisation at minimal additional cost.

8. The danger of derogatory use of collections, or their use implying an unsuitable endorsement

Concerns about derogatory use are met by all Creative Commons licences for the England and Wales jurisdiction (licences for Scots law are still being created).75 However, concerns about endorsement are not met by Creative Commons licences. The Creative Archive licence76 does include such a condition; but it does not restrict alteration of digital objects.77

9. Desire for attribution

All Creative Commons licences require attribution.

10. Retaining the ability to situate collections behind technical protection, and/or to time-limit or geographically limit their use (certain users only)

The issue regarding Creative Commons licences that seems most worrying is the fact that they are irrevocable.78 However, this simply reflects the inescapable fact that, once an image has been released onto the web, it is impossible to remove it definitively. The Creative Archive licence restricts use to the UK; but it does not restrict alteration of digital objects.79 Whilst at first sight Creative Commons licences seem to prohibit placing items behind technical protection, the CIE-CC report suggests that this is not actually the case, and that

---

[73] Barker et al., p. 31.

[74] See the 'case study' in Barker et al., p. 20.

[75] Barker et al., p. 17.

[76] Used by the BBC, Channel 4, Open University, British Film Institute and others: http://creativearchive.bbc.co.uk.

[77] Barker et al., p. 17.

[78] Barker et al., p. 12.

[79] Barker et al., p. 17.

‘Creative Commons licences do not restrict institutions from placing materials behind a password protection scheme’.80

Thus, as far as rights issues are concerned, the discussions held by the team responsible for the CIE-CC report suggest that, for low-resolution images, Creative Commons licences are likely to prove attractive. The most popular form of licence was BY-NC-ND (attribution, non-commercial, no modification), followed closely by BY-NC-SA (attribution, non-commercial, share-alike – i.e. implicitly allowing modification on condition of circulating the modified work under the same licence).81

11.9.9 Conclusions

Image provision from the museum sector is top-down, and consultation with communities of users regarding selection, formats, and metadata has not been extensive. The target audiences are usually an educated general public. Images are intended for personal and educational use, although restrictions are placed on their dissemination.

If museum collections are going to be incorporated into CLICs, individual communities will need to decide upon which images they require, in what formats and with what accompanying metadata. In addition, museums must be willing to meet these requirements.

There is currently no established mechanism for museums and communities to communicate with each other. This can be rectified by the creation of a two-stage communication mechanism, based upon a community/museum liaison committee and a museum digitisation committee.

Museums are likely to be concerned about a number of factors:

1. Losing the intellectual context for their images and data
2. Maintaining the most up-to-date version of digital objects, and potential degradations in quality as they are reused
3. The possibility of the unrestricted recirculation of digital objects
4. Allowing digital objects (particularly images) to be altered (certain owners only)
5. The difficulty of clearing third-party copyrights in parts of their collections to allow for further dissemination, and of identifying rights holders for ‘orphan works’
6. Potential loss of opportunities to generate revenue from their collections
7. The resources required to make collections available
8. The danger of derogatory use of collections, or their use implying and unsuitable endorsement
9. Desire for attribution
10. Retaining the ability to situate collections behind technical protection, and/or to time-limit their use (certain users only)

However, adoption of three basic guidelines should go a long way to alleviating museums’ anxieties:

4. Rather than create a single repository for digital images, aggregate data from individual image owners’ collections
5. Adopt standard licences which address these concerns (most likely Creative Commons licences and one or two newly-created variations)

---

[80] Barker et al., Appendices pp. 44-7.

[81] Barker et al., pp. 23-4 & Appendices p. 11.

6. Provide funding to assist in the exposure of existing collections and the clearing of third-party rights in those collections

11.9.10 References

Ed Barker, Charles Duncan, Andres Guadamuz, Jordan Hatcher and Charlotte Waelde, The Common Information Environment and Creative Commons: Final Report to the Common Information Environment Members of a study on the applicability of Creative Commons Licences, 10 October 2005 http://www.common-info.org.uk/docs/CC-Appendices.pdf (consulted 14 November 2005). The work itself, and therefore these extracts from it, are licensed under the Creative Commons Attribution Licence, http://creativecommons.org/licences/by/2.0/uk/.

The Digital Picture, December 2005, http://thedigitalpicture.ac.uk/documents/pdf/digital_picture_final_report.pdf (consulted 5 December 2005).

Digitisation in the UK: The case for a UK framework, Version 1.1, JISC & CURL: November 2005, http://www.jisc.ac.uk/uploaded_documents/JISC-Digi-in-UK-v1-final.pdf (consulted 24 November 2005).

Sites examined

Collection Name	Collection URL
Aberdeen Art Gallery and Museums: Explorer	http://www.aagm.co.uk/code/emuseum.asp
Applause South West	http://www.applausesw.org.uk/
Art and Architecture (The Courtauld Institute of Art)	http://www.artandarchitecture.org.uk/
Beamish Collections Online	http://www.beamishcollectionsonline.co.uk/
Bolton and Bury Treasures in Trust	http://www.ourtreasures.org/
British Library	http://www.bl.uk/onlinegallery/
Collage	http://collage.cityoflondon.gov.uk/
COMPASS (The British Museum)	http://www.thebritishmuseum.ac.uk/compass/
Cotton Town	http://www.cottontown.org/
Diduknow.info	http://www.diduknow.info/
Dorset Coast Digital Archive	http://www.dcda.org.uk/
Dulwich Picture Gallery	http://www.dulwichpicturegallery.org.uk/collection/search/dtlquery.as
Eadweard Muybridge Bequest	http://213.48.46.171/museum/muybridge/
EESOP Essex	http://eesopessex.essexcc.gov.uk/

Collection Name	Collection URL
Epact: Scientific Instruments of Medieval and Renaissance Europe	http://www.mhs.ox.ac.uk/epact/index.htm
Exploring the Potteries	http://www.exploringthepotteries.org.uk/
FACET: The many sides of life in Kensington & Chelsea	http://www.londonfacet.org.uk/facet/index.jsp
Fine Art Museums of San Francisco	http://www.thinker.org/fam/about/imagebase/index.asp
Habitas Online	http://www.habitas.org.uk/
Hamilton Palace - A Virtual Reconstruction	http://www.rcahms.gov.uk/hamilton/
Huntley & Palmers Collection	http://www.sopse.org.uk/ixbin/hixclient.exe?a=file&p=Huntley&f=huntley%2ehtm
I Dig Sheffield	http://www.idigsheffield.org.uk/
IMAGINE	http://www.imagine.org.uk/
Imperial War Museum: Concise Art Collection	http://vads.ahds.ac.uk/collections/IWM.html
Imperial War Museum: Posters of Conflict	http://vads.ahds.ac.uk/collections/IWMPC.html
Imperial War Museum: Spanish Civil War Poster Collection	http://vads.ahds.ac.uk/collections/IWMSCW.html
Ingenious	http://www.ingenious.org.uk/
Ingenious (National Museum of Science and Industry)	http://www.ingenious.org.uk/See/
International Dunhuang Project	http://idp.bl.uk/
Into the Net	http://www.intothenet.org/
Ironbridge Gorge Museums Online	http://www.ironbridge.org.uk/index.asp
IWM Collections Online: Photographs	http://collections.iwm.org.uk/server/show/nav.00g005
Knitting Together	http://www.knittingtogether.org.uk/

Collection Name	Collection URL
Moving Here	http://www.movinghere.org.uk/
MSIM: Virtual Collections Centre	http://82.138.231.51/web/pages/msim/collections/index.html
Musée du Louvre	http://cartelfr.louvre.fr/cartelfr/visite?srv=crt_frm_rs&langue=fr&initCritere=true, arts-graphiques.louvre.fr/fo/visite?srv=home
Museum of Modern Art, New York	http://www.moma.org/collection/search.php
National Gallery of Art, Washington DC	http://www.nga.gov/search/index.shtm
National Gallery, London	http://www.nationalgallery.org.uk/collection/default_online.htm
National Portrait Gallery	http://www.npg.org.uk/live/collect.asp
Natural History Museum	http://www.nhm.ac.uk/nature-online/online-ex/index.html,
Nature Navigator	http://www.nhm.ac.uk/nature-online/biodiversity/nature-navigator/
PeoplePlay UK	http://www.peopleplayuk.org.uk/
People's History Museum Online	http://www.nmlhweb.org/index.html
Petrie Collection	http://www.petrie.ucl.ac.uk/
PhotoLondon	http://www.photolondon.org.uk/
PortCities: Bristol	http://www.discoveringbristol.org.uk/
PortCities: Hartlepool	http://portcities.hartlepool.gov.uk/
PortCities: Liverpool	http://www.mersey-gateway.org.uk/
PortCities: London	http://www.portcities.org.uk/london/
PortCities: Southampton	http://www.portcities.org.uk/Southampton
Revolutionary Players	http://www.revolutionaryplayers.org.uk/
Rijksmuseum	http://www.rijksmuseum.nl/collectie/ontdekdecollectie?lang=en
Scottish Cultural Resources Area Network (SCRAN)	http://www.scran.ac.uk/
Shetland Museums Service Online	http://www.shetland-museum.org.uk/
Smithsonian Institution	http://smithsonianimages.si.edu/siphoto/siphoto.portal
Somerset Timeline	http://www.somersettimeline.org.uk/

Collection Name	Collection URL
State Hermitage Museum	http://www.hermitagemuseum.org/ fcgi-bin/db2www/browse.mac/category?selLang=English
Statistical Accounts for Scotland, 1799 and 1832	http://edina.ac.uk/statacc/
Tate	http://www.tate.org.uk/servlet/BrowseGroup?cgroupid=999999956
Ten Generations	http://www.tengenerations.org.uk/10Gen/index.jsp
The Metropolitan Museum of Art	http://www.metmuseum.org/Works_of_Art/
Victoria & Albert Museum	http://www.vam.ac.uk/collections/
Welcome to Powys: A Day in the Life	http://a-day-in-the-life.powys.org.uk/
Wilfred Thesiger Web Gallery	http://www.prm.ox.ac.uk/ThesigerWeb/
Window on Wiltshire's Heritage	http://www.wowheritage.org.uk/
Windows on Warwickshire	http://www.windowsonwarwickshire.org.uk/

11.10 Commercial Sector Review

CLIC Report from TASI

Report on Work Package 10: Commercial Collections

Document Notes

Author	Karla Youngs and Grant Young
Date	3 December 2005
Version	Final
Document Name	WP10-Report-final.doc

Contents

1. Introduction

This report covers WP10 undertaken by TASI. It is based on a desk-based investigation of nearly 70 image collections, most of which are listed or discussed in this report. Collections were identified from various sources, including TASI's Images Sites database and the membership list of the British Association of Picture Libraries and Agencies (BAPLA). Some collections were also identified through TASI's work on other CLIC work packages. Our focus has been on collections outside of education.

TASI's specific brief for Work Package 10 was to:

> ...examine the role played by image libraries, digital image sharing sites and electronic image archives in the commercial sphere ...[to] summarise the commercial sector's approach to metadata, rights management and access control, and see whether any useful operating models emerge. (CLIC Proposal)

We have taken a broad definition of commercial collections. Although we have concentrated on sites whose primary purpose is to sell images, we have also included some image collections that have built commercial activities *around* their collections (e.g. selling prints) or that are selling *services* rather than images (e.g. selling space for people to display their images). The work package brief suggested we should also look at sharing sites, so we have considered some collections that are non-commercial and based on sharing or exchanging images rather than selling them.

In looking for models or features that may be useful for CLIC's community, we have concentrated on six areas:

- Usability, interface design
- Metadata, search and retrieval
- Control of access and delivery
- Rights management
- Charging and costing models
- Community aspects

Each of these areas is discussed in detail in section 2 of this report. We pick out important themes or features, providing examples of each. Examples are drawn from a deliberately wide range of collections: some large and institutional; others very small and personal. URLs for all the collections mentioned are provided at the end of this report.

2. Features of commercial collections

2.1 Usability, interface design

It is in the interest of those selling images to make it clear what the scope and quality of their collection is, enabling those with particular needs to get very quickly to relevant images, and providing uncertain users with ideas and options. The interface plays a key role in achieving these ends. Its styling (i.e. look and feel) will differ according to the nature of the collection and its primary clients, but whatever its content and audience, it must be efficient in its navigation and effective in its communication.

Access to images

In reviewing the commercial sites, TASI has seen many that place access barriers at the very beginning (e.g. login forms) and provide few clues about the scope of the collection. An extreme example is the Garden and Wildlife Matters Photo Library. Apart from the title and 9 small images across the top of the page, no information is available unless a user registers. Although there is a separate Web site with some information about the Photo Library (http://www.gmpix.com/), this is not linked to from the online image collection. A similar example is DERWeb, a collection of dental images.

In contrast, there are sites that provide a good sense of the scope of their collection and lead the user very quickly into a selection of images. Art Resource, which provides access to images from museum and gallery collections, uses every opportunity to present the viewer with images from its collection. These are continually changing and are on most of its pages. When an image is clicked, the user is presented with a detailed record with hyperlinked keywords leading on to other images. In this way, users are led effortlessly through the collection and in the process gain a sense of its quality and scope. It is interesting to compare this with the Bridgeman Art Library, which has a similar collection and audience. Bridgeman provides a single image on its front page, which does not link anywhere. In order to get into the collection itself, the user must enter a search.

Usability and accessibility

Some collections adopt a very pared-down look, others are very busy in their styling. See for example Agripicture, which has a single picture and few words, or the Lightbox or UK Landscape Stock Photo Library, which have considerable amounts of text (which may be an attempt to boost their search engine ratings). However, a busy interface does not necessarily mean that the user will be frustrated in their attempts to access images. The Lightbox interface is confusing, but almost any link on the Landscape Stock Photo Library homepage will take the user straight to images.

A well-styled look is not necessarily a guarantee of good usability and accessibility. Dr Stock is delivered entirely via a Flash interface. Although the navigation is generally very effective, the user can have no control over the way they view the site. Untitled Picture Library is visually very interesting, but there are legibility issues with the choice of text and colours and the navigation cues can be misleading (e.g. some hyperlinks are underlined, some are not; there are prominent arrows which the user might expect to be highlighting links, but serve simply to link blocks of text with their

headings). More conventional navigation (e.g. top-tab or side-bar navigation) is generally more effective because most Web users will be familiar with its layout and metaphors (e.g. Crash Picture Association or Science Photo Library).

Information about the collection

It is clearly in the interest of an image collection to ensure its users understand how to use the resource. This includes help with the search, navigation and other technical aspects of the site. This information is typically accessed by via "help" or "FAQ" link (e.g. Mirror Pix, or the Medical Photo Library). Some collections provide extensive help in using their Web site (e.g. Getty Images); others provide little or no assistance. WireImage Stock, for example, instructs its users to phone a US number or email them with any questions about using their Web site.

In addition to technical help, commercial collections need to provide clear information relating to the purchase and use of images. This includes terms and conditions,. rights-related information and pricing. For some commercial collections, this is including among the help information. For others it is separate and often more obvious (e.g. Education Photos has a clear "How to Buy" link). We noted two innovative approaches to providing help or information about the collection. iStockphoto provides a simple Flash-based 'movie' explaining how their website and service works. Robert Harding Picture Library enables its users to initiate a "live help chat" with one of their image specialists.

Special functionality

Search and retrieval functionality is important for any digital image collection and is addressed in the next section. Another common and valuable asset for a commercial collection is a 'lightbox' feature. Most picture libraries use the term 'lightbox', since this metaphor comes from the photographic industry, but occasionally other terms are used, such as 'favourites'. Image After calls their lightbox 'clippings' and uses a scissors icon.

Typically a user must register in order to create a lightbox – and this is often a way of encouraging people to sign-up. However some sites enable their users to set up temporary lightboxes without registration (e.g. Bridgeman Art Library – for 30 minutes, or Dennis Kunkel Science Stock Photography – for the duration of the visit to the site). Others enable more 'permanent' lightboxes without requiring any registration (e.g. Photo Researchers Stock Photography, Image After). These non-registered lightboxes probably make use of 'cookies' (small text files) located on the user's computer – although the information must also be stored on the collection's server, since the user is often able to "email the lightbox" (generally a link) to themselves or to someone else (e.g. Image after or Photo Researchers Stock Photography).

Other functionality found in some collections includes community-related services such as forums and blogs. These are discussed in section 2.6 below.

2.2 Metadata, search and retrieval

The collections we looked at varied considerably in their use of metadata and their approach to search and retrieval.

Metadata

Most commercial image collections include some metadata, even if it is just a list of associated keywords (e.g. Photos.com). Others provide captions or information about the photographer, model (e.g. model release number), or image (e.g. size). Many include rights statements or credits. Unsurprisingly, collections that concentrate on providing stock imagery for use in design work typically provide less metadata than more specialist image collections (e.g. the Medical Photographic Library or Heritage Partnership collection - where we saw 200-word+ descriptions).

Most collections make use of structured metadata (i.e. metadata broken down into categories or elements – e.g. Bridgeman Art Library or Corbis). Mirrorpix, however, relies on a single field, which sometimes includes extensive information, and sometimes very little information or context. This has consequences for retrieval – and for understanding and using the images. We noted that sometimes the metadata categorisation itself can be confusing or misleading. The Heritage Image Partnership, for example, lists a "Creator" for its historical images. On investigation, this is the creator of the digital image, not the original image it reproduces.

The more structured the metadata, the more sophisticated a search or browse can be constructed. Because Mirrorpix has no categorisation to its metadata, it is only able to offer a simple search box (no advanced search or browse). This is not very efficient for a collection of over 250,000 images. Compare this with the Corbis advanced search, which is able to offer many different search parameters (including dates, orientation and colour) because this information is entered in a highly structured way.

Image searching and browsing

Often, collections can only be accessed via a simple search requiring the user to guess at the metadata terms that have been used. This is likely to frustrate and deter users. The cmpimages site is an example of a single search box site. It offers no advanced search or browse options, and few clues about the way the images have been catalogued. When the user gets to an image record, they can see some associated keywords, but cannot click on these to retrieve other images. Instead, they must return to the search box and enter another term. Dkimages and Garden and Wildlife Matters Photo Library provide further examples of very limited retrieval interfaces.

More effective search interfaces will provide clues about how the data should be entered (e.g. some suggested search terms), and may even enable the user to select terms from a list. Crash PA, for example, enables its users to build up a search by combining some simple drop-down choices. The search interface for Art Resource operates in a similar, but more sophisticated way. If a user enters a term in the "advanced criteria" or "assisted keyword" part of its advanced search form and then hits 'search', a drop-down list with all the relevant options is generated.

TASI found a very wide range of advanced search interfaces, from the fairly basic, perhaps enabling the user to switch from a Boolean AND to a Boolean OR or select/deselect a few filters (e.g. WireImageStock), to the highly sophisticated. We will describe some of the more interesting search features we observed.

- *Disambiguation/clarification*. Some of the commercial sites will ask users to clarify the meaning of ambiguous terms (e.g. Corbis, Getty Images or Inmagine). If a user types in a word that could have multiple meanings an intermediate page will ask them to select from a list of possible senses.

- *Search bias*. While it is fairly common for advanced searches to enable users to filter or prioritise their search by selecting or deselecting parameters, Fotosearch invites its users to choose a "search bias" (Asian, European, American or Latino). A default bias is set according to the version of Fotosearch begin used (e.g. the UK version has a default European bias).

- *Content-based retrieval*. Content-based retrieval draws on the perceptual content of the image (e.g. distribution of colour or shape within the image). These are important dimensions for those seeking images for graphic design purposes, so some stock photo collections will offer content-based searching. Some collections, such as Untitled, seem to rely on cataloguer-entered keywords (e.g. 'red', 'blue') to enable a simple content-based retrieval, but others, such as iStockphoto or Fotolia, appear to be using true Content-based image retrieval (CBIR) techniques, based on computational analysis. IStockphoto's search (http://www.istockphoto.com/file_search.php) enables its users to select colours from several different colour spaces, indicate which region of the image they want this to appear, and whether they want "tight" or "loose" conformity to these selections. Photos.com has a colour wheel for picking colours.

- *Theme-based retrieval*. In addition to supporting perceptual searching, commercial collections often support thematic searching, since designers often want images to illustrate a 'mood'. The Untitled search enables users to select 'emotional' keywords (e.g. "pessimistic", "fragile" or "bold"). Corbis indexes emotional terms among its keywords, but does not allow a separate search/browse on these.

- *"Inspiration"*. Some commercial collection provide features to help those who aren't quite sure what they're looking for. One approach is to offer a random selection. Photokey provides a very prominent "random search" from its homepage, while ImageAfter provides random images as one of its search's 'extras'. Bridgeman Art Library has an 'inspiration' section with examples of how its images can be used within graphic projects. iStockphoto has a similar set of articles. Christie's Images has an 'Ideas' section – which shows how its images might be used with words and other design elements.

In addition to searching, many collections enable their users to browse for images by subject category or by other image attributes. The ILN Picture Library offers broad thematic sections as a way in to its collection (e.g. "Fashion", "Sport"). Within these sections, users can view selected themes, browse all images, or search within the theme. Africa Photos does not have any search, but arranges its images into photo essays. FreeFoto.com, Buy Image, iStockphoto, and the Science Photo Library are further examples of collections offering subject category-based browsing.

Other common forms of browse offered by commercial collections include selections of the 'lastest images' (e.g. Crash Picture Agency), 'most popular images' (e.g. ArenaPAL, Big Stock Photo), or images/categories that are thought to be topical (e.g. Eyevine). Some collections with a geographical focus lead the user in via a hierarchical placename browse (e.g. View Buildings, or StockPhotography.co.uk). StockPhotography.co.uk also includes a map browse.

Sometimes browse features are integrated into the search interface – so that instead of typing in words, the user is able to rely on drop-downs of category-based filters. We have already mentioned Crash Picture Agency and Art Resources' assisted searches (above), but further examples are Photos To Go, which enables the user to paste keywords into the advanced search, and the Science Photo Library, which enables users to search a keyword dictionary and then select words to search on. The Science and Society Picture Library also provides access via its thesaurus. Unfortunately it does not give an indication of how many images are in each category and we found several terms that had no associated images.

Having retrieved results, it is fairly common for commercial collections to display the associated keywords and enable users to search/browse to similar images (although there are many that don't exploit this opportunity, e.g. Bridgeman Art Library). There are three main approaches to linking from results: (a) a simple 'find similar images' link, in which the system provides associated images based on its own criteria (e.g. Fotolia, Nucleus); (b) hyperlinking from keywords, which is in effect a browse, enabling the user to link to other images that have that individual keyword (e.g. Jupiter Images, Photos To Go); and (c) ticking a selection of keywords and then initiating a search (e.g. Robert Harding Picture Library, Corbis or Inmagine). The Heritage Image Partnership enables its broad categories to be hyperlinked at the item level, but not it's keywords, which it presents "for reference purposes only".

Commercial image collections offer other ways to order or refine the results of a search or browse. One of the more sophisticated we saw was Fotolia, which enables its users to refine the results further by colour information, image type (photo, illustration), orientation (horizontal, vertical), or license type. Results can be ordered by pertinence (presumably based on their own relevance algorithm), price, date or popularity. Fotolia also enables the user to email themselves the search criteria, so they can re-perform the search at a later date.

2.3 Control of access and delivery

It is in the interest of commercial collections to get the users into the images and demonstrate the value of their collection as soon as possible. So there is a balance between needing to show images and needing to protect images from misuse. Generally a commercial collection's approach is to enable users to get some way into the collection and view results, but to limit further access or deliverables or functionality, or to watermark their assets. Typically, certain functionality, like a lightbox is tied to registration and is a carrot for getting users to sign up.

Registration

Some collections place barriers at the very beginning. We have already mentioned Garden and Wildlife Matters Photo Library and DERweb. Others include Lonely Planet Images and Education Photos online. There are different approaches to vetting users. With DERweb, users are granted immediate access to the collection upon entering their details and access is later suspended if the email proves to be false (presumably if their email bounces back). With Wildlife matters, the registration is checked and then passwords are emailed.

The owner of Education Photos online phoned TASI up and asked several questions about the organisation. This may have been because we looked like an odd user, but is probably also due to the sensitive nature of the imagery – the picture library owner explained that he had to be careful where images of children were involved. Mediscan also have sensitive images. They advise people under 18 or who may be offended not

to enter the site. They also set three levels of access, providing level one access by default. Users must ask to be upgraded to other levels.

The other collections we examined enabled users to get some way into the collection before requiring registration (or all the way if it was giving away images for free or lacked any e-commerce facilities). Collections vary considerably in where they require registration or personal details. Epic Scotland enables its users to conduct a search, but will only display the results when an email address is entered (although this could be any address). Art Resource, Mediscan, Stockxpert, and many others require registration for a lightbox, while SeaPics enables users to set up and save a lightbox without any registration (it presents the users with a URL for bookmarking). Jupiter Images enables users to use a lightbox and a purchase 'cart' without any registration, but users must sign up to use their price calculator or to make a purchase.

Registration is often required to access larger images. Christie's Images is extreme, in not allowing access to anything beyond a thumbnail without registration. Dr Stock allows its users to look at five large preview images before requiring registration. A more common approach is to provide a watermarked preview image to everyone and enable unwatermarked and/or larger versions to those who have registered (e.g. Construction Photography, FLPA, Inmagine, Corbis or Getty).

Some form of registration is always required when e-commerce facilities are made available (e.g. Robert Harding Picture Library or Alamy) – although sometimes you can get almost to the 'check-out' without providing any personal information (e.g. DK Images).

In additional to registration, we noted that some sites asked users to click on an Agree/Disagree button to agree to terms and conditions (e.g. Visuals Unlimited) before being presented with images.

Image deliverables

Typically users of commercial image collections are presented with a thumbnail image which links to a larger preview image which is visibly watermarked. Increasingly collections seem to be providing an additional mouse-over enlargement to their thumbnails (see, for e.g., Image After, iStockPhoto, Fotolia). Jupiter images offers a thumbnail with mouseover enlargement, a preview image, and an even larger image for comping (i.e. laying out graphics in a mock up). Some collections enable their users to zoom on a portion of their preview image. Fotolia does this very effectively; Stockxpert, in contrast, only shows its users an enlargement of the middle of the image, which may not be the portion they wish to examine in detail.

Purchased images are delivered to users in many different ways. This may depend on whether the entire process is conducted online or whether there is offline negotiation, purchase and delivery. It also often depends on the size and format of the image required.

For some of the larger collections (e.g. Corbis and Getty) everything can be done online if the user wants to. In addition to selling individual images, Corbis and Getty also sell royalty free CDs, and will provide the purchaser with online access to their 'virtual CD' (Corbis for a year; Getty, indefinitely). Some collections provide the options of online downloads or posted CD-ROMs (e.g. Adams Picture Library); some have certain images available for immediate download and others that will require later delivery (e.g. DK Images). The approach taken by Science and Society is to enable images to be purchased online, and zipped preview images to be downloaded for comping, but to email the high resolution images. Others (e.g. Agripicture) will conduct their sales offline, but place the images onto a server for their clients to download. In general, images that are made available to download or provided via

email are large, compressed JPEGs. If clients want TIFF images, these are typically put into the post on a CD-ROM.

'Open' collections

Generally collections offering free images place few or no barriers on access and enable easy downloading of their images (e.g. Morgue File, Image After). However OpenPhoto, which uses Creative Commons licenses, now requires its users to register in order to access and download its larger images. Registration is automatic (an email confirmation is sent).

2.4 Rights management

There are several different approaches used to protecting rights within a digital environment. Some are *educational* – telling people what they can and can't do with the resources. Others are based on managing users or deliverables, for example *limiting access* via systems of authentication and authorisation, or *limiting deliverables* by providing lower resolution versions online or stamping images with a visible watermark. There are also some higher-end approaches involving file sealing and encryption.

Limiting access has been discussed in the previous section, along with the limitation of deliverables (e.g. access to larger images). We will concentrate here on educational approaches to managing rights and on the use of digital watermarking. It may be that some of these collections are using high-end technologies in delivering the higher resolution images to clients. TASI was unable to observe this, since we did not purchase images in the course of our investigation.

Educational approaches

The simplest approach to rights management is to tell users about the copyright situation. Some make no explicit mention of copyright (e.g. Natural History Pictures); others rely on generic copyright statements (e.g. CMP Images, Jupiter Images). Sometimes there is a link to a separate page detailing copyright. Art Resource provides an article explaining copyright, while Agripicture provides links to other Web sites which deal with copyright (US sites). Sometimes information about copyright is found under the Web site's 'Terms and Conditions' (e.g. Alamy, Christie's Images).

Some commercial collections place image-specific copyright information alongside each image (e.g. Heritage Image Partnership or DK Images). Bridgeman Art Library provides links to rights holders for its in-copyright images (search on Picasso, for an example), while Photo Researchers Stock Photography provides a downloadable text file with all the copyright and credit details for each image.

Among the free image sites we looked at, Open Photo makes use of Creative Commons licenses with its images, while Image After has terms and conditions that are its own, but very similar to Creative Commons.

Licensing within the stock photo industry takes two main forms: *rights managed*, where images are licensed for specific, restricted uses, and *royalty free*, where images are purchased and can be used many times in a variety of ways. These are discussed in section 2.5 in relation to charging models.

Watermarking

Visible watermarking is a very common feature of commercial image sites. Some commercial collections do not watermark their images (e.g. Art Resource or Untitled Picture Library – but note that the preview images they provide are very small). Others watermark some of their images (e.g. ArenaPAL and Stockxpert). In the case of ArenaPAL the variation is because the collection is drawing its images from different sources. Stockxpert makes a small proportion of its images available for free use – so it does not watermark these. Some collections add very discreet watermarks (e.g. Medical Photographic Library). Alamy, puts its information along the bottom of the image in a way that doesn't mar it. Others have extremely strong watermarks that obscure information in the image (e.g. Eyevine, Heritage Image Partnership). Some watermarks contain more than just a logo or the name of the collection. FLPA, for example, includes an image reference and a credit statement. The Science and Society Picture Library includes the number of the particular image and it also dynamically stamps on details of the user (username, or guest, and date of access).

Invisible watermarking is by definition more difficult to observe. We are aware that Corbis uses the Digimarc service (http://www.digimarc.com/) and it is likely that others do as well. Unlike visible watermarking, invisible marking does not limit the use of the image. It is only useful as a way of proving ownership and – if expensive tracking services are paid for – detecting infringement.

In addition to visible watermarking, other general approaches to restricting use are to wrap the file in another format (e.g. Dr Stock's use of the Flash format), using a zoom function to show part of the image (e.g. Fotolia) and disabling the right-click or copy functionality of the browser. These approaches to limiting access to larger images are very commonly employed within cultural heritage settings, but are much less common with commercial collections. The reason is that commercial clients generally want to enable their users to create mock-ups or 'comps' before purchasing the images. Restricting copying of the preview image would frustrate these users. For commercial collections the visible watermark is a preferred and more effective approach.

2.5 Charging and costing models

Stock photo collections have traditionally offered two main models for pricing their collections: *rights managed* or *royalty free*. Rights managed images are available for specific uses (e.g. one placement in a book or displayed in an advertising campaign for a limited time period) and are typically more expensive and exclusive than other options. Royalty free images are cheaper and have fewer restrictions. Once you've purchased a royalty free image you can use it again and again.

In addition to these two main models we have observed and describe three other models: *subscription-based* collections, *credit-based* collections and *free* collections. Subscription collections enable users to download large royalty free images for a particular time period. Credit/exchange collections enable users to build up credit by either purchasing it with money, or by contributing and selling their own images. Free collections offer images that are free for many purposes, although they usually come with some restrictions and sometimes fees are charged for commercial applications.

Rights-managed and royalty-free collections

Some collections are entirely rights-managed (e.g. DK images, Education Picture Linrary, FLPA); others are entirely royalty-free (e.g. Inmagine); others offer both (e.g. Getty or Corbis). Royalty free images are offered at flat rates, usually based on the size of the image. For example, all the images the Adams Picture Library sell are offered at the following flat rates: £49 for A6 printing and Web; £129 for A5 printing; and £199 for A4 printing (all plus VAT). Although its images are royalty free, Adams is willing to turn many of its images into rights-managed images for a negotiated fee. A client might want to do this where they want to guarantee that they are the only one using that image.

Flat-rate fees are easy to calculate and work well within an e-commerce environment. Rights-managed pricing is more difficult since it is based on a number of different factors. Many commercial collections negotiate offline over their rights managed images (e.g. Art Resource or Natural History Pictures). Some provide indicative pricing online (e.g. Science Photo Library), but many provide no price guidance at all (e.g. Untitled Picture Library). Where rights-managed collections want to make use of e-commerce, they must provide online price calculators. Examples of these include Heritage Image Partnership, Mediscan, and the Robert Harding Picture Library. We saw others, but many are only available upon registration (e.g. Alamy, Construction Photography).

We noted that some collections will explicitly licence rights-managed images for educational use (e.g. teaching). Mediscan charges £8.84 for a standard resolution (i.e. screen-based) image for academic use (unspecified). Nucleus Medical Art will provide an image for PowerPoint for one year for $65, while the Welcome Trust's Medical Photographic Library sells a digital image for PowerPoint for £2.94 for an unlimited duration (compared with £5.88 for their slides of the same images).

Other models

Increasingly those offering royalty-free images seem to be turning to *subscription-based* models of delivery. Inmagic offers subscription access to some of its collections. Photos to Go provides rights-managed images from $39 and royalty-free for a flat $79 per image, or unlimited access to its royalty free images via subscription ($49 for 1 month up to $149 for 12). Photo-key and Photos.com are entirely subscription-based. Prices vary considerably, but so do the sizes of the collections. Photo-key, for example, charges 79 euros per month or 249 euros per year to access "thousands" (low thousands) of images, while Photos.com charges $139 per month or $599 per year to access 170,000 images. Although these collections say you pay one fee and then download all you need, there are charges for the very highest resolution images, and there are usually limits to the number you can download within a 24 hour period (50 for Photo-key; 250 for Photos.com).

Credit-based collections operate on a kind of exchange model. The collections are built up through user contributions, with contributors receiving a credit for each of their images that is sold. They can cash their credits in or use them to buy other people's images. Those who just want to buy images purchase credits online with a credit card and then use these credits to buy the images they want. The system is based on royalty-free images, with cheap flat rates per image (sometimes graded by size). iStockphoto is probably the largest collection operating on this model. It claims to have 350,000 members and 529,000 images. Contributors earn $0.20 for every $1.00 credit spent on their images. These can be cashed in once they reach $100, or can be converted to 'download credits' to spend on other people's work. New contributors must submit images for checking before they become approved contributors. Once a contributor is approved, each image submitted is still checked individually for file size, quality and copyright. A more recent entrant is Stockxpert,

which has grown out of a free image collection called Stock.xchng. Stockxpert is much more generous in its royalty payments to contributors than iStockphoto, giving contributors 50% of the purchase price and cash payment once $50 is reached. A very interesting feature of Stockxpert is that it enables users to commission work. Users are able to set up a project and then invite image contributors to provide a quote for the work.

Fotolia also operates on a credit-based system, with a much more complex pricing and remuneration structure. In addition to standard royalty free fees (based on size) sellers earn more if they contribute their images solely to Fotolia, and if they offer an exclusivity buy-out license (i.e. enabling buyers to take up exclusive usage rights). Under these different criteria, a contributor to Fotolia can earn anything from 33-80% of the purchase price. The buyer might pay anything from 1 credit (£0.57) for a Web-sized image up 1500 credits (£855) for an exclusivity buy-out. Another credit-based collection is Big Stock Photo.

We looked at several *free image collections*, including Image After, Open Photo, Morgue File, FreeFoto.com, and FreeImages.co.uk. Some are stocked with images by individuals or small groups of people (e.g. Image After, FreeImages.co.uk). These tend to be quite small in size (e.g. FreeImages.co.uk has 2,500 images, although we note that FreeFoto.com has 76,000 images). Other free image collections are open to contributors (e.g. Morgue File or Stock.xchng), but are generally vetted for quality. Stock.xchng has a basic criteria for contribution and a further system of certification which marks out those contributing the best quality images. Some free collections can be genuinely used for any purpose (including commercial) without payment or permission (e.g. Morgue File). Others are free for personal use, but will charge for commercial use (e.g. FreeFoto.com).

We also looked at two large, general image sharing sites: Flickr and Webshots. Neither vet for quality, as the ones in the previous paragraphs do. They are primarily sites for people to showcase their own personal images, although Webshots has built some e-commerce around their site, mixing in stock photographs (in the Gallery section) and enabling users to buy prints of their photographs or pay for "premium" services. Flickr offers a "pro" level of service and carries some advertising, but has otherwise avoided commercialisation. Those contributing images to Flickr are able to attach creative commons licenses, giving permission for certain uses. Flickr is also notable for its user-generated metadata, and its community aspects (discussed in the next section).

In addition to the models describe above, we note that there are services that might be described as providing *portfolio services* for photographers or artists. Eyevine looks like a unified stock photo collection, but is in fact a "digital railroad" service providing both a showcase for photographer's work and a download service. The actual negotiation over price and the payment is, however, arranged directly between the photographer and client. Photographers pay Eyevine a set-up fee, and subscriptions for services and storage space. UK-based Artindustri enables artists to promote and sell their physical works. It includes portfolios for 2,500 artists with about 22,000 images online. Artists can create a free portfolio, which allows them up to 20 images, an artist statement, contact details and links and the ability to generate e-cards from their images. For $99 they can upload as many images with as much text, to multiple galleries. Artindustri does not offer any e-commerce features: users and artist must deal directly over purchases or commissions.

2.6 Community aspects

Finally we looked at community aspects of commercial and shared image collections. Unsurprisingly, collections that are free or encourage user contributions offer a more social experience, although it is clear that community and commercial elements can go together. Fotolita describes itself as a "social marketplace for royalty free stock images". IStockPhoto calls itself a "royalty free stock photography community."

We will describe some of the community or social features we observed from the collections examined.

Personal identity

Some collections give their contributors an identity. In the more commercial stock photo collections this might take the form of showcasing a professional photographer's work (e.g. Robert Harding Picture Library), but in the shared, contributory collections, this often takes a more personal form. For example, the Morgue File has a page for each of its contributors which includes some basic demographic information (location, interests, gender), links to photographs they've uploaded, any comments or guestbook entries that have been made about their work and any comments they've made on other people's work. iStockphoto goes further, enabling its member/contributors to add a photograph of themselves, biographical information, and have their own blog (see, for example, http://www.istockphoto.com/user_view.php?id=177123). Flickr and Stock.xchng are further examples of collections that enable their contributors to create a personal identity.

Ratings, reviews

Many commercial collections enable their users to view their most popular images, based on the number of downloads (e.g. Big Stock Photo). Others make this more personal by inviting users to rate or review other people's images. iStockphoto enables both reviews and ratings. Although it does not offer a search filter based on ratings (an omission), it does enable users to view the "most popular files". These are arranged in two sequences: one according to ratings given over the previous three months; the other according to downloads.

Other communication features

We found many features designed to support communication, particularly among the shared or credit-based collections. These include more static features such as guestbooks, reviews and comments (e.g. Open Photo), and more interactive services, such as forums (e.g. iStockPhoto or Stock.xpert) or blogs (FreeFoto.com, iStockphoto). Some also have wiki's (e.g. iStockPhoto has a technical wiki). Open Photo originally took a very open approach to its collection, enabling users to edit metadata and contribute to the wiki. Because of abuse, much of this has now been restricted. When we tested it, we were able to edit the categories and our own photos, but not anyone elses. Flickr provides many tools for connecting people and images, including comments, testimonials, and RSS feeds. Eyevine, which is a hosting/promotion service rather than a unitary collection (described above), also provides RSS photo feeds as a way of promoting images and drawing people into the collection.

3. Conclusion

This report has covered a wide range of features of commercial and free image collections, many of which will be relevant to community-led image collections. We found some good examples of design, metadata use, access control, and rights management, along with some examples of poor practices or missed opportunities. There are useful lessons for CLIC in each.

List of collections

This section provides links to the collections discussed in the body of this report.

Adams Picture Library – http://www.adamspicturelibrary.com/

Africa Photos – http://www.africaphotos.com/

Agripicture - http://www.agripicture.com/

Alamy – http://www.alamy.com/

ArenaPAL – http://www.arenapal.com/

Art Resource – http://www.artres.com/

Artindustri – http://www.artindustri.com/

Big Stock Photo - http://www.bigstockphoto.com/

Bridgeman Art Library – http://www.bridgeman.co.uk/

cmpimages – http://www.cmpimages.com/

Chiro.org Image Archive – http://www.chiro.org/chimages/

Crash Picture Agency – http://www.crashpa.net/

Christie's Images – http://www.christiesimages.com/

Dennis Kunkel Science Stock Photography – http://www.denniskunkel.com/

DK Images – http://www.dkimages.com/

Dr Stock – http://www.doctorstock.com/

ePicscotland – http://www.epicscotland.com/

Eyevine - http://www1.eyevinearchive.com/

FLPA – http://www.flpa-images.co.uk/

Fotolia – http://www.fotolia.co.uk/

Fotosearch – http://www.fotosearch.co.uk/

FreeFoto.com – http://www.freefoto.com/

Garden and Wildlife Matters Photo Library – http://www.gardenmatters.uk.com/

Heritage Image Partnership – http://www.heritage-images.com/

ILN Picture Library – http://www.ilnpictures.co.uk/

Image After – http://www.imageafter.com/

Inmagine – http://www.inmagine.com/

iStockphoto – http://www.istockphoto.com/

Jupiter Images – http://www.jupiterimages.com/

Lightbox – http://www.the-lightbox.com/

Lonely Planet Images – http://www.lonelyplanetimages.com

Medical Photographic Library (MPL) - http://medphoto.wellcome.ac.uk/

Mediscan – http://www.mediscan.co.uk/

MirrorPix – http://www.mirrorpix.com/

Morgue File – http://www.morguefile.com/

Natural History Pictures – http://www.natural-history-pictures.co.uk/

Nucleus – http://catalog.nucleusinc.com/

Open Photo – http://openphoto.net/

Photo-key – http://www.photo-key.com/

Photo Researchers Stock Photography – http://www.photoresearchers.com/

Photos.com – http://www.photos.com/en/

Photos To Go – http://www.photostogo.com/

Robert Harding Picture Library – http://www.robertharding.com/

Science and Society Picture Library – http://www.scienceandsociety.co.uk/

Science Photo Library – http://www.sciencephoto.com/index.html

SeaPics – http://www.seapics.com/

Stock.xchng – http://www.sxc.hu/

Stockxpert – http://www.stockxpert.com/

UK Landscape Stock Photo Library – http://www.buyimage.co.uk/

UK Stock Images – http://www.ukstockimages.com/

Untitled Picture Library – http://www.untitled.co.uk/

View Buildings – http://www.viewbuildings.com/

Visuals Unlimited – http://www.visualsunlimited.com/