



## Visualising China – final article / report

Project name	<i>Visualising China, ILRT, University of Bristol</i>
Project website	<a href="http://visualisingchina.net/">http://visualisingchina.net/</a>
Report author	<i>Nicky Ferguson</i>
Project period	<i>April 2010 – March 2011</i>



### Synopsis

*Visualising China* has created a web site bringing together a number of important collections of photographs of China from 1850 to 1950. These collections, some already existing digitally and some newly digitised, are cross-searched and browsed from within one interface using a *Linked Data* approach. In addition several social software features have been included, allowing users to comment, make suggestions on specific attributes of the images, such as date, location, people and places depicted and their significance. It attempts to combine the flexibility of crowd sourcing and informal annotation with the structure and provenance of archived collections.

The design and the software engineering have both been heavily influenced by user testing and usability work conducted at different stages during the project.

We have successfully incorporated more images and more collections than we anticipated in the project plan. Our priorities on features and functions have shifted several times during the project in response to user feedback. The project itself has given us the opportunity both to build a diverse user group and to meet with other developers on related JISC projects and exchange technologies, approaches, opinions and insights.

We have also brought together an impressive list of future collaborators who have agreed to make further substantial and very important image collections available through the integrated Visualising China interface, should future funding be forthcoming.

## Background to the project and context

1. Visualising China is an online resource<sup>1</sup>, developed over the last 14 months as a research project with JISC funding, that significantly enhances access to a number of important collections of photographs and metadata relating to pre-1950's China. As well as giving access to many individuals' albums from trunks and attics, it includes Bristol's digital collection of nearly 7000 historical photographs<sup>2</sup> integrated visually with the complementary collections held at Virtual Shanghai<sup>3</sup> and Queen's University Belfast's Sir Robert Hart Digital Image Gallery<sup>4</sup>. Visualising China currently offers an extended cross-search, cross-browse and novel visualisation of the collections and related resources in a Web 2.0 context, introducing potential valuable related China eContent from Google Books. It will allow scholars, as well as those interested in family history and the history of trade and relations between China and the UK, to annotate and comment upon images as well as include their own collections.

The screenshot shows the 'visualising CHINA' website interface. At the top, it says 'China 1870–1950: an interactive resource' and 'Logged in as Jonanna Blogs | Sign out'. There is a search bar and a 'My workbench (34)' button. The main content area displays a photograph of a man sitting in a chair in a room. Below the photo, there are several metadata fields: 'Location' (San Ya, Guangdong province), 'Date' (1920-1930), 'Keywords' (Fu Bingchang, Chinese government organs - 中國政府部門, Politician), and 'Notes' (University of Bristol - Historical Photographs of China reference number: Ar02-009). There is also a 'Map' showing the location in China and a 'Related images' section with several small thumbnail photos. At the bottom, there is a comment section with fields for Name, Email address, and Comment, and a 'Submit' button.

attics, it includes Bristol's digital collection of nearly 7000 historical photographs<sup>2</sup> integrated visually with the complementary collections held at Virtual Shanghai<sup>3</sup> and Queen's University Belfast's Sir Robert Hart Digital Image Gallery<sup>4</sup>. Visualising China currently offers an extended cross-search, cross-browse and novel visualisation of the collections and related resources in a Web 2.0 context, introducing potential valuable related China eContent from Google Books. It will allow scholars, as well as those interested in family history and the history of trade and relations between China and the UK, to annotate and comment upon images as well as include their own collections.

2. The "Historical Photos of China" digitisation effort has taken place over the last five years within the 'Chinese Maritime Customs Service' project based in the Department of Historical Studies at the University of Bristol. Between 1842 and 1954 tens of thousands of Britons visited or lived and worked in China. Bristol's digital collection offers access to thousands of photographs of life in China before 1950, donated by families or descendants of Chinese and British men and women who formerly lived there. These images include snapshots of expatriate social life, scenes of everyday Chinese life, and records of momentous political events. They are a valuable resource for the study of Chinese modern history.

3. Part of the "Historical Photos of China" open access collection<sup>5</sup> is currently hosted on the French collaborator's server. In addition, Visualising China cross-searches and browses two major UK digital collections building an interface from elements fetched from diverse international sources. Visualising China has developed a reliable and long-term sustainability strategy based on a hosting solution with the full support of the Advanced Computing Research Centre at the University of Bristol.

<sup>1</sup> In the final development stages, see: <http://visualisingchina.net/>

<sup>2</sup> The total Bristol collection is 15,000 digital images with approximately 7,000 being "website-ready" with full metadata mark up, expected to grow to more than 10,000 in the next 12 months via an ongoing digitisation process separate to this project undertaken as a continuation of the 'Chinese Maritime Customs Project'. Original funding by the AHRC, the Chiang Ching-kuo Foundation for International Scholarly Exchange, the 1999 Higher Education Funding Council Chinese studies initiative, through the East Asian Institute at the University of Cambridge, the Universities' China Committee, John Swire & Sons Limited, The British Academy, WUN, and the University of Bristol.

<sup>3</sup> <http://virtualshanghai.ish-lyon.cnrs.fr/>, a research and resource platform on the history of Shanghai from the mid-19C

<sup>4</sup> <http://digitalcollections.qub.ac.uk/>

<sup>5</sup> Approximately 5000 digitised photographs are included in the Virtual Shanghai collection,

4. The Visualising China software offers:

- **Cross-search and browse of the photographs**, through a cross-mapping of metadata from the Historical Photographs of China, the substantial Sir Robert Hart collections at Queen's University Belfast (Sir Robert Hart was Inspector General of the Imperial Customs, Peking, 1863-1908, and a key figure in China's 19th century history and its foreign relations with the West<sup>6</sup>), Virtual Shanghai, the most recently added Joseph Needham Photographs of Wartime China, 1942-1946 from the Needham Research Institute, Cambridge and other collections.
- **Image annotation and general social software support** for personal or shared use by researchers. This allows researchers to make specific, digital cross-references between people, locations, images, time periods and related resources (such as books, film archives, newspaper excerpts) in an extensible manner.
- **Support for further cross-linking** to user annotations of early 20<sup>th</sup> Century China newsreels and photographs uploaded to Flickr. This will allow for (moderated) informal, as well as formally collected multimedia content to be linked in to Visualising China.
- **Visualisation of the images and related content** through interchangeable context views (such as by timelines and via location with views of geographical maps as well as more traditional text-based views).

This project has already excited interest amongst varied communities of researchers interested in visual history and modern China (researchers in Environmental Change, Politics, Economics and Law, as well as History). Originally developed for the Screen Arts community the underlying software approach not only illustrates the benefits for Chinese modern history research, but will be documented, disseminated and freely available for further use throughout the JISC community and beyond.

## The issues the project is addressing, why the work is important and to whom

### Dispersed, scattered and hard to find resources

Researchers and those teaching increasingly popular courses in Chinese Studies find it hard to bring together and discover images of China from the period 1850-1950. There are a number of reasons for this, one being the number of social upheavals during that period in China itself with many photographs being lost or destroyed; another being the fact that many important photographs are as yet undigitised, some still in trunks or attics in the UK and all over the world; another being the difficulty in accurately and narrowly searching for relevant digitised images on the web (as opposed to text which, with the advent of powerful indexing and search engines, is much easier to track down). Most of the well-known image searches such as Picasa, Flickr, Google Images and PhotoBucket will present a large amount of irrelevant material to even a sophisticated searcher.



### Information on many images is patchy at best

Many images come to us from albums kept by people whose ancestors were in China as merchants, military personnel or diplomats. They may or may not be annotated and the annotations may or may not be meaningful to us ("*Chalky and Bones on shore leave*"). Many thousands of miles and perhaps a hundred years intervene between the images and the locations they depict. Even with the best kept, meticulously recorded collections, key information may be missing and the items shown in the photographs which are of interest to today's historians (or family historians) may not be the ones that interested the photographer or cataloguer at the time.

<sup>6</sup> Queens University holds 77 volumes of Hart's personal diaries (1867-1908), covering his whole career in China, notes and papers relating to Hart's experiences and analysis of the Boxer disturbances in Peking, 1900, plus numerous photographs, slides etc of Sir Robert, his relatives, friends and colleagues, c 1854-1911

## Addressing the issues

The original proposal included our intention to "build on sophisticated visualisation Web software from the JISC-funded STARS project. We will upload the images to a Bristol-hosted database searchable by STARS software and extend STARS technology to support image annotation". So the first tranche of effort was spent developing the technology to bring images from existing collections and allow web clients to display them in the VC interface – for this we used a custom-built version of the STARS software which would allow us to accomplish the following:

- Rapidly include newly digitised materials and other collections
- Search and browse them in intuitive ways
- Allow a new community of users to add to and update the information
- Also allow them to add new images of their own from personal collections on Flickr or Picasa or newly digitised images from hard copies or indeed pointing to resources already available on the web.

We felt it was vital to use a designer and a user experience expert to ensure that the interface we were building would be useful to the target community.

We were gratified to find that a community of intending and enthusiastic users coalesced rapidly around our prototyping work and continue to help us engage with our target audiences. The user engagement workshops have proved invaluable, helping to emphasise for our team the importance of solving information problems as well as solving technical problems; holding a small focussed technical workshop also worked well, allowing us to share ideas with projects clearly working in the same area and with insights to contribute. The feedback from the user experience work, however, led us to transform the user interface of the STARS software so that what had previously been a side-panel at the edge of graphical visualisation instead became the core of the user experience, with visualisations being called in as required by context. This, combined with extended tagging, commenting and annotation interfaces built in the latest Google Web Toolkit (GWT), has added a wealth of Web 2.0 features to the software foundation carried through from STARS.

## Challenges, achievements and lessons learnt

### 1. Research project versus user-driven utility

Many of the challenges we met during the project fell under this heading. The project was funded by JISC to investigate and prove the use of social software solutions for academic purposes – sharing historical images of China for teaching and research. However the funding is not for "pure research" – the intention is to provide an example of a web site which uses the technology, offers a utility to an identified community of users (and perhaps others) and is indeed used by that community. We studiously avoid using the "s" word – "service", since it implies an ongoing commitment to providing for the users; but of course the niceties of this will be lost on most of the web site's users. So the project, like many others funded in this way, has to reconcile twin aims of proving a technical approach within a limited amount of time and funding and yet providing a web site which we hope people will use for some years afterwards (with all the implications that has for support, amendments, updates etc). Not an easy circle to square.

**Achievement:** creating a web site whose content and features appeal to its target audience.

**Lesson:** clarify the role and longevity of deliverables and the (lack of) support likely to be available when research funding ends.



2. **Pros and cons of incorporating design and user-experience work in a research project**

Designer Ben Hayes and user experience expert Stuart Church were involved in the project from the outset. There is no doubt that their presence and their input provided us with a totally different end product from what would have appeared without their involvement. Because we had this involvement, we were perhaps over-confident as the project neared its conclusion – "we do not anticipate major changes as a result of the user testing process" was one phrase which in retrospect tempted fate. In fact, major changes were exactly what were required and we had to change the underlying technology (and in turn rapidly redesign to incorporate the requested inclusion of new functions and omissions of old ones) in order to deliver these changes. So the user testing transformed not only the interface but also the technology and the approach.

**Achievement:** a usable and attractive web site featuring the functions which its target audience require and uncluttered by those features they neither understand nor value

**Lesson:** even when design and user experience is incorporated in the project at an early stage, the crucial test is when the technologists deliver enough functionality to put a working prototype in front of the users – expect significant changes at that stage.

**Home: Click on the home page elements you like**

The screenshot shows the homepage of 'Visualising China: China 1870-1950: an interactive resource'. It features a search bar, a user login area for 'Joanna Bloggs', and several content sections: 'Bringing Chinese History to Life' with a list of image collections, 'Most popular tags' word cloud, and 'Latest activity' feed. Colored circles (hotspots) are overlaid on the interface to indicate where users frequently clicked or interacted. Notable hotspots are located on the search bar, the 'Latest activity' section, and the 'Most popular tags' word cloud.

- Search box is obvious and on the first page
- Good to be able to feel like part of a community working with the photographs
- Uncluttered design
- More interesting than tags, I think [Latest activity]
- Nice to have this at the start of the site [Collection list]

Part of a 14 page usability report on a candidate interface for *Visualising China*. The Colours denote "hotspots" i.e. places where most users clicked and user comments are seen at the bottom of the figure.

### 3. Challenges and requirements for using social software in an academic setting

While this project and many others are keen to bring the advantages of social networking to learning, teaching and research, academic enterprises (and particularly the institutions which typically host them) have concerns and imperatives which are often different from both commercial and informal social collaborative sites. We highlight a few issues here:

#### a. Authentication

For ease of use, we decided to use OpenID as the main authentication mechanism for the site. This means that anyone with an account at Yahoo, Google, Orange, MySpace, Flickr and several other providers can sign in to Visualising China without creating a new ID and password. While this is neat and convenient for users it proved in user workshops to have two disadvantages: first (and surprisingly) many academic users, particularly older or more established ones, either did not have or could not remember the details of any of these accounts; secondly it is very easy to create an account with one of these providers without accurately and fully identifying yourself, so where controversial material is being handled and commented upon, this may cause a problem for the institution responsible for the site. Which brings us on to the next issue:

#### b. Provenance

One of the aims of the project has been to bring together resources from multiple collections into one familiar interface. Of course it is important for users to know where the image and description(s) have come from. Even for images whose ultimate provenance is in doubt (i.e. exactly where and what date they were taken may be unknown), the fact that they form part of xyz collection is an important piece of information.

Equally, it is important for users of a research/learning resource to know whose comments they are reading and where the images they are looking at have come from.

*Skywalker* or *Happymum* may be fine for blogs or YouTube but they will not help someone putting together a pack of referenced learning materials or a research proposal.

An academic focussed site has a particular challenge in balancing such considerations against the highly desirable ambition of avoiding bureaucratic barriers to use. We also faced the challenge of judging "what is core data?" and "who can change it?" Even sites such as Wikipedia wrestle with this issue, with most sections open to change from any user (with a fallback facility) but an increasing number of important or controversial sections needing approval for changes.



#### c. IPR, copyright, legal and policy implications

At an early stage we formulated a policy for dealing with IPR which included a Creative Commons statement<sup>7</sup> and negotiating appropriate use agreements with a variety of rights holders. However the introduction of social software techniques brings a series of new challenges. In controversial areas, allowing users to comment and add or suggest information will need an element of oversight. Clearly a site with academic pretensions would not be keen to host legally dubious claims of whatever sort. Allowing users to "add" new resources from the Internet (even if those resources are not physically saved by Visualising China but merely pointed to "as if" they were part of a physical collection) also raises questions of rights and ownership. Users must be made aware of their

<sup>7</sup> This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported Licence](https://creativecommons.org/licenses/by-nc-sa/3.0/). Based on a work at [www.visualisingchina.net](http://www.visualisingchina.net). For higher resolution images or permissions beyond the scope of this licence contact us by email: [hums-chinaphotos@bristol.ac.uk](mailto:hums-chinaphotos@bristol.ac.uk)

responsibilities and requested to assert that they own the rights or have permission to make these images available through Visualising China. A simple and rapid "objections" procedure also needs to be in place allowing copyright holders to object and an administrator to take down any alleged offending material while an investigation is made. Similarly objections should be allowed to comments or annotations.

d. **Combining metadata**

A related problem is that of dealing with and displaying a wide variety of metadata. For each of the formal collections we are including in the interface, formal catalogued metadata (usually added recently) for each image/resource is available. These are provided in a variety of schemes (usually home-made) and of varying quality within and between the collections. In addition we have a variety of comments made much earlier: some are written on photographs, some on albums and some on the negatives or slides themselves; some were added at the time the photograph was taken, some much later in another country; some were made by the photographer and some by cataloguers or the subjects or their descendants. Add to this the tags, annotations and suggestions that users will be adding to the site. Many questions arise from this: do we treat them all the same – as one body of interesting information? ... do we search all the comments/metadata in the same way but attempt to differentiate the sources in the search results? ... do we rank sources in different ways? ... do we add to this melange by making available extra information on the "collection", the photographer, the copyright holder, the collector? Lots of interesting questions for which we don't pretend we have found all the answers. In addition, we have the task of making a bridge between the terminology and jargon of the collector/librarian/archivist and the jargon of modern web and social software users - how do *collections*, *classification*, *keywords*, sit alongside *tags*, *tweets* and *like me*? Anyone undertaking a similar task will need to be aware that combining comprehensive, scholarly information with the clarity and simplicity needed for a usable web site is a fascinating and complex challenge.

**Achievement:** We have created a prototype which we feel integrates many of the attractive and useful features of social networking and crowd sourcing while addressing issues of provenance and integrity which are important to academic researchers.

**Lesson:** The imperatives and priorities of an academic website may be different from those of a commercial or informal service. It remains to be seen whether the audiences feel or act differently. We anticipate many more lessons once the site is widely used.

4. **Sustainability – technical and content**

"Staying alive" is an important issue for a web site once short term funding for development ceases. JISC have recognised this and required institutions to put in place sustainability measures to ensure that the site does not disappear shortly after the research funding ceases. But these measures tend to be focussed on the technical side – ensuring that server space is found, URLs are persistent and that scripts continue to work when no-one is looking after them full time. However to a user there are other equally important "signs of life". These entail changes to the home page and enough up to date references to appear cared for. References to a "recent meeting" held a number of years ago would not be signs of life! So sustainability would require at least some attention to content and appearance – of course where the resource comes from to fund that continuing attention is another matter.



**Achievement:** We have found and nurtured a mutually advantageous relationship with the ACRC, mentioned above, which gives us the technical sustainability we and JISC require.

**Lesson:** To sustain "signs of life" and freshness, a website must have some level of continuing input – ideally from its users.

## 5. Angels and Mentors

We were very fortunate to have input from various outside experts and advisors. In particular, we welcomed visits from different JISC staff, our experienced and supportive advisory group, including Malcolm Raggett from SOAS, and our project mentor Grant Young, Digitisation and Digital Preservation Specialist from Cambridge University Library. While most projects have involvement from senior academic figures as an institutional contribution, it is rare to have such active and enthusiastic participation as we did from Professor Robert Bickers. The participation of such angels and mentors in a project may sometimes seem inconvenient and challenging (especially when they are suggesting more work!) but we have found it to be always constructive and thought provoking.

## 6. Workshops and user evaluation

Our workshops, both user-focussed and technical, proved to be a source of great inspiration and ideas. In a practical session with users who have never been exposed to the interface before, especially where the audience is very varied, in seniority, language, technical expertise and historical speciality, things happen which no-one expects. For us these were both positive and negative, with users identifying things to do with our data which we had not considered and others failing to notice or use features which we thought had been made simple and obvious. In particular, our user experience consultant was able to conduct detailed interviews independently of the team in an environment where frank criticism and detailed discussion of workflow and task orientation were encouraged. The results were most useful in the iterative design and construction of the site.

## Impact

The project made an immediate impact within the community of academic historians interested in Chinese history of the period because we were fortunate enough to be able to demonstrate it at several appropriate events and keep the "early adopters" informed by involving them in testing and evaluation. It is uncertain exactly how use of the site will evolve, but the fact that several more prestigious institutions have expressed interest in making their important collections available through Visualising China augurs well for the future. In addition, the project's technical work and the workshops it led created much



interest in the community of developers interested in creating social software solutions for academic and research information sharing. Simon Price presented a paper at the ninth UK e-Science All Hands Meeting (AHM Sept 2010) <http://www.allhands.org.uk/>

Within the institution, the project has also made an impact much bigger than its size. This is mainly because as part of our sustainability plan, we were fortunate enough to have the cooperation of the new University of Bristol Advanced Computing Research Centre<sup>8</sup>. At the well-attended and high profile recent launch of the ACRC, Visualising China was heavily featured<sup>9</sup> and the launch was also prominently covered by regional media.

Selections of the images featured in Visualising China have been exhibited at the Brunei Gallery, the Museum of East Asian Art, Bath, the Oriental Museum, Durham University, Grant Bradley Gallery, Bristol, Red Navarra de Estudios Chinos (Navarre Network of Chinese Studies), and the Universidad Pública de Navarra (Navarre Public University), as well as being shown in other exhibitions in Beijing, Nanjing and Hong Kong. 'Picturing China 1870-1950: Photographs from British Collections' by Robert Bickers, Catherine

<sup>8</sup> <https://www.acrc.bris.ac.uk/>

<sup>9</sup> <http://www.bris.ac.uk/news/2011/7446.html>

Ladds, Jamie Carstairs and Yee Wah Foo was the subject of a 12 page supplement to International Institute of Asian Studies Newsletter and the HPC collection was also featured in Articles in Asian Art, Research Horizons (University of Cambridge), Bath Life, Swire News, Evening Post and two University of Bristol publications: *Subtext* and *Research* .

Images were also featured in the University of Bristol centenary celebrations/ celebration of Chinese culture, and 'China: Journey to the East' at the City Museum & Art Gallery).

While as yet we have scanty information on Visualising China's impact on practice in teaching, learning and research, we expect it to build on these high profile activities and feedback on potential use from user testing has been very positive.

In the longer term we hope that the project will bring together academic historians, commercial companies (we already include the Swire collection from John Swire and Sons Ltd<sup>10</sup>), those researching family histories and those still alive who have personal memories or archives of the key people and places featured in the site.

## Conclusions and plans for the future

The project started as a plan to use an existing software platform to bring more images of China to academics and to provide those users with more functions and capabilities. It turned into a fully fledged software development project, bringing social networking together with academic and commercial historical archives.

With an adaptable and collaborative team and the help and support of many outsiders we believe we have produced a good example of how we can begin to bridge the gap between formal academic research needs and the increasing familiarity of users with simple, easy to use interfaces and the capabilities of social software.

Visualising China will be formally launched to the public at the international conference *Treaty Ports in Modern China* 7th-8th July 2011 and the British Inter-university China Centre's conference on *Britain and China, pasts, presents and futures, from the nineteenth century to the twenty-first*, 24th-26th August 2011.



In the future, we plan to develop the web site, bringing in further collections from The British Library, The Needham<sup>11</sup> Research Institute, Cambridge, The School of Oriental and African Studies (SOAS) and the Bath Royal Literary & Scientific Institution who have all indicated their enthusiasm. We also plan to tackle the metadata challenges created by integrating new collection and item level information from disparate content collections by developing lightweight tools to aid and integrate both the formal cataloguing of images as well as informal, crowd-sourced metadata "tagging". We would like to build new user communities, both through the use of social media and by contact with recent Chinese alumni and elderly members of the Chinese Diaspora. We would also like to offer a tool from our JISC-funded MyMobileBristol project to allow smartphone users in China to augment image data with geo-locations and also to take their own photos of these locations to help build up a "then and now" view of these depictions of life in China. Although the funding landscape and priorities are quite unclear at this difficult time for the UK economy, given the new priorities of the Arts & Humanities Research Council (e.g. 'Digital Transformations') and also the current JISC review, it seems likely that some emphasis will move away from "blue sky" towards collaborative and immediately productive research work. This project, with many collaborators from different sectors and combining the input of archivists, commercial institutions, artists, historians and technologists and a clearly important international dimension might be seen as a model for such future work. With China an increasingly significant economic superpower and Chinese History entering A-level and GCSE syllabuses, the development and exposure of these rare resources is relevant not only to researchers but potentially to all those involved in trade and relations with China who wish to understand better China's recent cultural and political history.

<sup>10</sup> [http://en.wikipedia.org/wiki/Swire\\_Group](http://en.wikipedia.org/wiki/Swire_Group)

<sup>11</sup> In fact, we are very grateful to the Needham Institute for supplying us with their data and allowing access to their images in advance and we have included them experimentally in our cross-search for this release, though more work needs to be done on fully incorporating the collection and integrating our approach to the metadata.