How Web 2.0 will change history

Possible futures for websites of the National Archives of Australia

Catherine Styles

In this paper I define (or at least characterise) Web 2.0; survey Web 2.0 technologies and applications that have a bearing on the historiographical life cycle of publishing, research and development; and introduce the National Archives of Australia as a

- publisher of historical material
- research facility, and
- capacity-builder of student historians

My ultimate question is: how could Archives' websites adapt to the Web 2.0 environment and culture, or ... where to now?

Defining Web 2.0

'Web 2.0' simply refers to the emergent generation of web tools and applications. It's a term with many detractors, for whom the term reeks of marketing hype and is devoid of real meaning. But it is now widely used, and significant, as Tim O'Reilly (of the O'Reilly publishing group, which popularised the term) suggests, when he says: 'Like many important concepts, Web 2.0 doesn't have a hard boundary, but rather, a gravitational core' (O'Reilly, 2005: 2).

The mind map of Web 2.0 that appears on Wikipedia and in many other places on the web shows key examples of Web 2.0 applications (for blogging, wikis, photo-sharing, e-commerce); concepts like user-centred, modularity, and collaboration; and key systems of the emergent web, such as social software and standardisation.

For my purposes, the key term is 'social software'. A huge part of what Web 2.0 does is facilitate community. To some extent, it's all about dating. But software can also build communities of professional practice. So, for example, you might use the photo-sharing site *Flickr* to join or set up a group of people interested in photographs of ruins. (I made that example up, but there is such a group – with 561 members.) Or you might use *citeulike* to organise and share your own bibliography of journal articles, and to leverage the power of other readers with similar interests.

Social software is software to which users contribute content, and therefore whose content gets richer, or more accurate, the more people use it. So you can begin to see how the professional (and personal) lives of historians might be enhanced by Web 2.0 applications.

As well as having fuzzy boundaries, Web 2.0 applications are proliferating at a great rate. There is an ever-growing list of Web 2.0 applications that currently runs to over 20 pages. When I was trying to get a sense of Web 2.0 as a whole, and how it might change the practice of producing history, I needed a sense of how Web 2.0 would impact specifically on the professional lives of historians. How historians use the web is obviously quite particular. I wanted a visualisation that was specifically tailored to the practices of producing history.

So I made my own 'mindmap' of Web 2.0 applications relevant to the life cycle of historical production – publishing, research and development. This diagram provided me with a framework for this paper.

Historiographic life cycle

My focus here is practices in which the National Archives is publicly engaged – as a publisher of histories (for example in its *Uncommon Lives* series of biographical stories), as a research facility (a provider of access to records of the Australian Government), and as a capacity-builder, for people learning to

research, develop and publish history, through its education resource *Vrroom* – virtual reading room.

The one item that had to go in the centre of the mindmap, because it is relevant to all three stages of historiography, is social software, which – because it is such a core part of Web 2.0 – I have already introduced.

Publishing

In this brief survey I begin at the top, with the publishing phase of the historiographic cycle, and how the evolution of Web 2.0 has changed the publication of history.

Wikis

There is a significant power shift occurring in the publishing industry. The most popular encyclopedia is no longer Britannica; it is Wikipedia (wikipedia.org). Wikipedia has a huge array of information, which is ever-growing, and ever-improving. The amazing thing about it is that anyone can edit any page. Every reader is a potential writer. Publishing no longer communicates one-way; it's all about participation.

The other amazing thing about Wikipedia is that it is also authoritative. You can argue that point with me, but according to a study published in *Nature*, it is as accurate as Britannica (cited in Ferris & Wilder, 2006).

This opening up and flattening out (and some might say dumbing down) of the practice of historical publishing can also be seen in sites like *The Remembering Site* (www.therememberingsite.org). If you can't leave your children money or land, leave them 'the most precious commodity of all – family treasures made of words'.

The State Library of Victoria (www.slv.vic.gov.au) is braver than many cultural institutions, in that it invites users to 'share what you know' about every digitised image in its collection. The link goes simply to a web form – there's no sense in which your response will appear on the site immediately. But it's a start.

The National Archives has a website called *Uncommon Lives* (uncommonlives.naa.gov.au) which is a series of historical biographies of individuals, couples and in the case of the next story, due out in April 2007, a community of Muslim Australians. *Uncommon Lives* is one site where it may be highly appropriate to host an adjunct wiki, so that people who know more about the subject can add detail, or contrast, or possibly even counter-claims. Such a feature would be really interesting as a complement to the perspective you'd get from the government records. (This is not to say that government records are always dry or impersonal; only that there will always be other perspectives.)

Podcasting

Another key feature of Web 2.0 is the broadcasting of audio and (increasingly) video, which users can download to consume either in iTunes or on their portable player. The Victoria & Albert Museum in London (www.vam.ac.uk) has a podcasting channel through which they publish – and to which you can subscribe – short discussions about various objects in their collection.

The National Archives of Australia is also experimenting with podcasting. We've been testing video and audio versions of a lecture on the fall of Saigon – given that it's the 30th anniversary of the fall of Saigon, we've also just released records on that topic.

We'll also probably podcast smaller, more easily digested chunks of audio, perhaps an audio-plus-images version of our popular 'Find of the Month' record and story. Find of the Month is present on the website but also in the gallery in the Archives in Canberra.

We'll probably also podcast stories relating to other objects on display in our new introductory gallery due to open in March 2007. The idea here is that a podcast about an object might inspire new audiences to visit the physical gallery.

The Victoria & Albert Museum adds value to its podcasting service by linking to other resources on the topic of the current podcast. For example, when they podcast about Che Guevara and design, they included a link to 'See what other people are writing about Che Guevara'. That link takes you to the Technorati search results page for Che Guevara. Because Technorati is an engine for searching the world of blogs and blog posts, this brings me to another large part of the landscape of Web 2.0 publishing – blogs.

Blogs

Blogs are an easy way to publish, and a great way to give and receive feedback on ideas before they are published in a more substantial form. They are becoming an adjunct to the academic journal system. So far I've not found a blog tool designed specifically for producing online academic journals, but I'm sure it's only a matter of time. Or maybe this is an indication of the lack of support among academics for this form of publishing?

An appropriate application of blog technology in the National Archives might be to host a blog for our Summer Scholarships program. Every year between one and three third-year or honours students come for six weeks to research and write about a topic in the collection. The Archives could set up a blog to publish their work-in-progress and the results.

RSS

The really interesting technology underpinning blogs and many other Web 2.0 applications is RSS – really simple syndication. RSS is a technology that helps you manage information overload, especially to keep up with the news. Many websites, including that of the National Archives, have an RSS feed for announcing new services, new features on their site, or news related to their content.

You can subscribe to the news, so that it appears in an abbreviated form through an RSS reader, along with all the other content to which you have subscribed. If it interests you, you can follow the link and see the item in full. Browsers like Firefox, Safari and Flock have a 'live bookmarks' feature that enables you to see the headlines directly through the browser. As an alternative to using live bookmarks in a browser, you can also use a dedicated tool for managing all your subscriptions. In Google's free reader, for example, you see the most recent posts on the left, with the source – the blog – named underneath. Bold ones are not yet read. The highlighted one is shown in preview form on the right, with a link to the full item in its original context.

The Archives is currently working on an initiative to publish a 'Pic of the week' – a visual record, whether a photograph or a plan or poster – on our new home page (due in early 2007). A feature that changes with that frequency will enable us to respond in good time to special events or anniversaries or seasons. It would also be an excellent stream to feed out to people to place on their own website, with a link back to our site. We would benefit from the increased attention, and they would benefit from having regular, interesting content that they didn't have to develop or maintain.

Radical trust

There is one aspect of the Web 2.0 landscape that is really significant for publishers, whether they are cultural institutions like archives or libraries or museums, or historians like yourselves. Web 2.0 demands a *radical trust*, on the part of publishers, of their users (Mazar, 2006).

Wikipedia would never have mobilised the minds of millions of people, it would simply not have worked, without the default principle that anyone can edit anything. People love to leave their mark, to *participate* in the production of knowledge, rather than to be told what something means without any opportunity to respond.

This issue of authority has been circling around and in some instances infiltrating cultural institutions for decades. Traditional paragons of authority, museums, galleries, libraries and archives will have to come to terms with participatory publishing in the next few years: in the emergent landscape of Web 2.0, the entire notion of authority and how you get it is changing. Institutions that publish without participation, that continue to rely on the

strength of their traditional authority, and which fail even to embrace a notion of shared authority, may find that their relevance and influence wanes.

The discipline of history is equally challenged by this shift in the process of authorisation. In that sense Web 2.0 can be seen as a new force in the history wars. And this is where my paper tries hardest to live up to its title, by going so far as to suggest that as the practices of reading and writing history change, so too may the history wars be tempered by this flowering of authorisation. Concrete documentation of a massacre would be irrefutable; but a powerful story is a powerful story.

Research

I'll turn now to look at the research stage of the historiographic life cycle, and how historians might benefit from Web 2.0 evolutions in searching: aggregation, visual browsing, smarter searching.

Aggregation

As the content available on the web has become richer and richer, so providers are getting together to produce portal sites, rather than expecting that users will discover relevant information by going to each of their sites one by one. Increasingly, resources are aggregated so that users can find and sort large, diverse sources of data through a single interface.

The National Archives' own *Australia's Prime Ministers* is an example of a site that serves as a gateway to primary and secondary sources. But the access it provides goes only to the top level of information in other collections. You can navigate through to where the collections are, but from there you have to search anew.

Picture Australia takes a more evolved approach, in that it serves you digital content from other sites. It does that by regularly harvesting metadata from all its partner sites, digging down into their collection and outputting thumbnails and descriptions of the records seamlessly in the *Picture Australia* interface. Then you can click through to the site from whence the data came, to see the record in its original context.

In the United Kingdom, The National Archives has just aggregated the holdings of 11 collections – held by The National Archives, local and private archives – into one search interface.

And *ArchiveGrid* is a United States-based site that aggregates data on archives around the world – hundreds of collections.

In the principle and practice of aggregating search facilities, we're getting closer to the ideal of the semantic web, where all web content is interoperable and machine-readable.

Of course, the National Archives of Australia's online collection interface, RecordSearch, remains exemplary in that it provides access to so much fully digitised content, and continues to provide a digitisation-on-demand service.

Visual browsing

Another feature of historical research in the Web 2.0 world is that sources are findable through multiple means. You don't have to think like an archivist or librarian, to navigate to what you want. You can use a visual map. Or you can browse a tag cloud.

A tag cloud can provide a simple way in to a collection, if you're not sure what's in there. The collection database interface of the Powerhouse Museum includes a tag cloud – it is generated according to what words users have keyed in to search the database.

For those unfamiliar with these things, tag clouds are created via an algorithm that sets the font size of the term according to its relative significance – so the most popular search terms appear largest.

On Amazon, tag clouds are used to represent a single resource – a book – as a way of finding out what's inside. The tag cloud shows the most frequently used words in a book. Again, the most common words appear largest.

Tag clouds are really just a way of outputting keywords that in the olden days would have stayed inside the database, as metadata, used in the background to help you find what you're looking for. Letting users see the keywords attached to a record, or to an entire collection, is good in principle, in that it renders visible a part of the process of making the records accessible. But it also serves a very practical purpose, in that users can get an overview, at a glance, of what's inside the collection, its most popular themes – its *character*.

From the point of opening up a collection in this way – allowing a user to see the keywords attached to a record – the next step is to enable them to contribute their own keywords. Like the concept of radical trust, this is a significant part of the Web 2.0 paradigm shift.

The Powerhouse Museum (www.powerhousemuseum.com) may be the first museum in the world to implement user tagging for describing its collection. For anyone who works with a very vast collection, the thought of a potentially endless supply of volunteers to help describe it must be appealing. The Archives, for example, could never afford to pay for the work that the collective interested public could perform.

It would also be very useful for the entire community of Archives users. Not only would they be empowered to leave their mark and to participate in interpreting the records, they would also find that the records became,

through this process, more findable. Connections would emerge between objects that may not otherwise be formally related.

If you think about a collection like the Archives, which is arranged by government function rather than by subject, you can begin to imagine how user-generated tagging could make a collection more usable, more meaningful, and more valuable.

The *Steve* project (steve.museum) is researching and experimenting with tagging in the context of art museums.

Smarter searching

Web 2.0 also facilitates smarter searching – you can take a recommendation from the server as to what might interest you. This used to be done, as it still is in *Vrroom*'s paper trails, by content developers nominating related records in the content management system. So the report on a second interview with a person applying to migrate to Australia would be linked to the initial interview report, and the correspondence that elicited the later interview.

But in the new Web 2.0 world order, such links are created by the application, as it tracks your choices and estimates relevant pathways. We've all seen Amazon's 'Customers who viewed x also viewed y'. Dutch archivist Eric Ketelaar called for Amazon-style recommendations on archives websites in 2003 (Ketelaar, 2003: 3).

Amazon, of course, has a suite of nifty tools for enhancing your search experience: search inside for tag citations; which books this book cites; and which books cite this book, and so on.

These kinds of advances have yet to appear on an archival website, but the Powerhouse Museum's new collection search facility, as well as allowing users to add tags to any object also, according to Sebastian Chan, Manager of Web Services,

> tracks and responds to user behaviour recommending 'similar' objects to increase serendipitous discovery and encouraging browsing of our collection. It also keeps track of searches and dynamically ranks search results based on actual user interactions. (Chan, 2006)

In 2007 the Archives is launching an entirely redeveloped website. In the next phase, we'll install a content broker so that we can set up the kind of personalised searching that is necessary to satisfy users in the Web 2.0 world. For example, you'll be able to customise what displays on the home page when you visit – eg, display forthcoming exhibitions and events but not the latest in records management.

A more significant step for historians will be when the collection database, RecordSearch, facilitates a smarter form of searching. Such an initiative will not come forth in the next year, but it will happen.

In the next few years we'll certainly see more and more of the collection described at item level, and more and more of the collection fully digitised. We may see a facility for users to tag records as they view them – a complementary folksonomic classification system to emerge alongside the archival classification system.

Perhaps as an alternative, or in addition, we'll use technology for generating tagclouds from smart scans of the text on our records. Server recommendations based on tracking your own searches and other similar searches would be nice. We'll almost certainly offer multiple RSS feeds so that you can track, for example, newly opened or newly digitised records (in the last x days, or since my last visit). And probably you'll see the National Archives collection aggregated in new ways with other collections. Our photographs are already accessible via Picture Australia. And we're a key partner on the *Australian ScreenOnline* website (the National Archives has ten times more audiovisual material than the National Film and Sound Archive).

Development

As well as enabling historians to publish their work in new ways, and to find both secondary and primary sources, Web 2.0 can facilitate the process of drafting new historical works. There are many, many tools out there to aid you in your work of organising (and sharing) data, and in collaborating with others to discuss, evaluate and review sources, and to draft, edit and produce history.

The one site I do want to mention here, not because it is an exemplary Web 2.0 application, but because it goes further than most primary source sites in the direction of Web 2.0, is *Vrroom*, virtual reading room for teachers and students (vrroom.naa.gov.au). The Research component enables you to annotate, sort and save your found records.

Historians' capacity-building

I want to turn briefly to consider the value of Web 2.0 tools for helping people learn how to do history.

There is plenty of scholarly debate about how Web 2.0 is being applied to teaching and learning, even in the traditional discipline of history. A lot of the debate centres on the value or otherwise of Wikipedia. In the history department at George Mason University, undergraduates and graduates alike are encouraged to contribute to wikipedia as part of their study. For Professor Mills Kelly, students use it anyway, so rather than ignore that, it's better to incorporate it into the course, and encourage its *critical* use – to raise

consciousness among students of its shortcomings and failings, and to join in improving it (Kelly, 2006).

Vrroom is an excellent tool for training the historians of the future and is headed in the direction of Web 2.0, even if it is not really there yet. *Vrroom* is a good laboratory for the Archives to experiment with new functionality, since it is a much smaller set of data, aimed at a less traditional and therefore more receptive audience (students and teachers).

I'm going to end with a few points about the future of *Vrroom:*

- It would be easier to implement a folksonomic system of keyword generation in *Vrroom* than it would be in RecordSearch. The debate has only just begun about the value of tagging versus traditional classification systems. So to implement it in the context of a discrete, and limited, set of records will be simpler both technologically and politically.

- *Vrroom* could easily incorporate an RSS feed for users to keep track of new content.

- The facility it already has for users to annotate and group their records could become socialised, ie we could make groups shareable.

- A more ambitious idea we have is to develop an online board game within *Vrroom,* which could be played alone or in groups, which would challenge you to find a record that is linked in some way to a previously played record. The nature of the link might be fixed according to the theme of the game, or it might be more open and dependent on the wile of the player. The aim would be to think, however deeply or critically, about the records individually and as a group.

To conclude, the web is becoming more useful for historians, easier to use, and more fun to use. And if institutions like the Archives want to remain relevant, they would do well to yield to – indeed to join – the force that is Web 2.0.

Dr Catherine Styles is Websites Editorial Manager at the National Archives of Australia. This paper was presented to the Australian Historical Association's conference in Canberra in July 2006.

References and further reading

Blog and listserv posts

Chan, Sebastian, post to the *Steve* discussion list – subscribe via steve.museum, June 2006

Kelly, Mills, 'Whither wiki', *Edwired: A weblog devoted to the teaching and learning of history online*, 14 December 2005, chnm.gmu.edu/history/faculty/Kelly/blogs/edwired

Mayaud, Christian, 'All things Web 2.0 – "the list"', *Sacred Cow Dung: Mythocracy in venture capital, technology, healthcare, media, internet et al*, 10 March 2006, www.sacredcowdung.com

Mazar, Rochelle, 'Radical trust', *Random Access Mazar*, 11 May 2006, www.mazar.ca

Morville, Peter, 'Authority', *Semantic Studios*, 11 October 2005, semanticstudios.com

Ratcliffe, Mitch, 'Wikiproject on the history of science', *Rational Rants*, 5 February 2006, blogs.zdnet.com/Ratcliffe

Shirky, Clay, 'Semi-structured metadata has a posse: A response to Gene Smith', *You're It: A blog on tagging*, 27 August, 2005, tagsonomy.com

Smith, Gene, 'Peter Morville: the tagsonomy interview', *You're It: A blog on tagging*, 19 October, 2005, tagsonomy.com

Walker, Jill, 'Network literacy: Learning with blogging and Web 2.0', *jill/txt* (jilltxt.net), 20 April 2006

'Web 2.0 is people', *Wink Blog: Notes from a start-up*, 10 March 2006, blog.wink.com

Papers

Alexander, Bryan, 'Web 2.0: A new wave of innovation for teaching and learning?', *Educause Review*, March/April 2006

Bearman, David and Jennifer Trant, 'Social terminology enhancement through vernacular engagement: Exploring collaborative annotation to encourage interaction with museum collections', *D-Lib Magazine*, vol. 11 no. 9, September 2005, www.dlib.org

Cameron, Fiona, 'Digital futures II: Museum collections, documentation, and shifting knowledge paradigms', *Collections: A Journal for Museum and Archives Professionals*, vol. 1, no. 3, February 2005

Cohen, Daniel and Roy Rosenzweig, 'Web of lies? Historical knowledge on the internet', *First Monday: Peer-Reviewed Journal on the Internet*, vol. 10, no. 12, 2005, www.firstmonday.org

Chun, Susan, Rich Cherry, Doug Hiwiller, Jennifer Trant, Bruce Wyman, 'Steve.museum: An ongoing experiment in social tagging, folksonomy and museums', *Museums and the Web 2006*, Albuquerque, 22–25 March 2006

Ketelaar, Eric, 'Being digital in people's archives', *Archives & Manuscripts*, vol. 31, no. 2, November 2003

Leadbeater, Charles, address to the session 'Libraries and the creative economy' of the symposium Library of the 21st Century, State Library of Victoria, 23 February 2006

O'Reilly, Tim, 'What is Web 2.0: Design patterns and business models for the next generation of software', O'Reilly Network, www.oreilly.com/1pt/a/6228

Sompel, Hervert van de, Sandy Payette, John Erickson, Carl Lagoze, Simon Warner, 'Rethinking scholarly communication', *D-Lib Magazine*, vol. 10, no. 9, September 2004, www.dlib.org

Spadaccini, Jim, 'Museums and the Web 2.0', CHIN Roundtable: E-Learning in Museums, powerpoint presentation, via www.ideum.com/blog

Weibet, Stuart L., 'Web Wise 2006: Inspiring discovery – unlocking collections' conference report, *D-Lib Magazine*, vol, 12, no. 3, March 2006, www.dlib.org