



# Archives 2.0

**- a summary of the way Archives NZ  
could use Web 2.0 technologies and  
approaches**

**Written by Julian Carver, Seradigm**

**23 May 2008**

**This paper is published with the permission of Archives NZ. It is shared in the hope that it will benefit archival institutions and other parties.**

**This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/>**

## *Table of Contents*

<a href="#">1 Introduction.....</a>	<a href="#">3</a>
<a href="#">2 What is Web 2.0.....</a>	<a href="#">3</a>
<a href="#">3 Web 2.0 Tools and Technologies.....</a>	<a href="#">4</a>
<a href="#">3.1 Blogs.....</a>	<a href="#">4</a>
<a href="#">3.2 Wikis.....</a>	<a href="#">5</a>
<a href="#">3.3 Tags .....</a>	<a href="#">6</a>
<a href="#">3.4 Social Bookmarking.....</a>	<a href="#">7</a>
<a href="#">3.5 Social Networking Services.....</a>	<a href="#">7</a>
<a href="#">4 Use at Archives NZ.....</a>	<a href="#">8</a>
<a href="#">4.1 External Use.....</a>	<a href="#">8</a>
<a href="#">4.1.1 By the Public.....</a>	<a href="#">8</a>
<a href="#">4.1.2 By Government Agencies.....</a>	<a href="#">9</a>
<a href="#">4.2 Internal Use.....</a>	<a href="#">9</a>
<a href="#">4.3 Risks.....</a>	<a href="#">10</a>
<a href="#">4.4 Culture Change.....</a>	<a href="#">10</a>
<a href="#">5 References and Further Reading.....</a>	<a href="#">11</a>

## 1 Introduction

Archives NZ is entering a new phase. A transition is taking place from the agency being focused primarily on print records to a future incorporating significant emphasis on digital content. Archives NZ will be taking responsibility for ‘digital born’ content, and will undertake a large programme of digitisation of old records. The agency will also act as a champion across the whole of government for digital record keeping, and good information management.

As such it must embrace new technologies which support digital information management. There are a number of new approaches enabling collaborative content creation, description and sharing that have emerged under the ‘Web 2.0’ banner. These include blogs, wikis, tags, social bookmarking, and social network services. This document provides a brief summary of these technologies and some recommendations on their use in the Archives NZ context.

## 2 What is Web 2.0

Web 2.0 is a phrase coined by O’Reilly Media in 2003, and made popular by the first Web 2.0 conference in 2004. It referred to a set of new technologies, methods and philosophies for doing business on the World Wide Web.

In the opening talk of the first Web 2.0 conference, Tim O’Reilly and John Battelle summarized what they saw as key principles of Web 2.0 applications. The most relevant of these to the archival context are:

- ***the web as a platform*** – that software doesn’t so much run on individual desktops or servers, it runs on the Internet. The web is a software development platform itself
- ***data as the driving force*** – that data created by users, and by user transactions is more valuable than data/content created by large institutions for passive consumption
- ***network effects created by an architecture of participation*** – the more the systems allow people to participate and interact, the more the system, its user community, and its value can grow, with that growth occurring in a non linear (typically exponential) fashion
- ***innovation in assembly of systems and sites composed by pulling together features from distributed, independent developers (a kind of "open source" development)*** – that software can be built using a combination of software components (called web services) running in real time, from a range of different web sites that expose their data and functionality for use by others

Web 2.0 is often contrasted with the ‘walled gardens’ of content and functionality of the first generation of Internet businesses. Much of the content in Web 2.0 sites is user generated, and can be shared dynamically between sites in real time.

Attitudinal traits of the Web 2.0 philosophy tend to include openness, inclusion, a tolerance for some disorder, valuing ‘amateur’ contribution, and a willingness to experiment and take adaptive/evolutionary rather than failsafe approaches.

### 3 Web 2.0 Tools and Technologies

Web 2.0 tools turn the 'read only' web to the 'read write' web by making it easy for users to contribute content. They include:

- Blogs
- Wikis
- Tags
- Social Bookmarking
- Social Networking Services

#### 3.1 Blogs

A Blog is a tool for a single or small group of authors to create and broadcast content to a (typically) large audience. The authors write content (text, images and hyperlinks) in a diary style format, called 'posts'. Each post has a date, title, and body. These are displayed in chronological order on a web site where the desired audience can read and comment. Blogs encourage a strong separation between author and audience participant. Audience members cannot modify published blog posts. They can, however, comment on posts, emulating perhaps a printed newspaper or magazine article with its accompanying letters to the editor.

One of the features that have made blogs very popular is syndication. Blog posts can be syndicated as a 'feed' to another web site, or to a software tool for reading blogs. In this way a user doesn't have to go back to a blog's web site to check for new articles, rather new posts can be 'fed' directly into a portal homepage or blog aggregator/reader application that the user look at regularly.

There are different types of blogging including individual, corporate, and internal. Regardless of their types all blogs have a distinctly personal 'voice' and are very different from polished corporate press releases. Their tone is often informal and exploratory, although it can equally be inflammatory and contentious as with some political blogs. Corporate blogging involves individuals in a company blogging to the world at large. These blogs are very often focused towards their customer and supplier communities. Both Sun Microsystems and Microsoft are using corporate blogging, often by very senior executives (the CEO in Sun's case) to powerful effect in creating dialogue with their customer base.

Although the idea of actively limiting your audience runs somewhat counter to the blog ethos, the blog is experiencing wide adoption within the corporate intranet environment. [Intranets Today](#) has published an in-depth article on the use of blogs within the corporate intranet, and the term 'internal blogging' is regularly used to describe blogs geared toward use in this domain.

There are many different software tools for blogging. There are hosted solutions, open source servers, and proprietary products. Popular hosted solutions include [Blogger](#) and [Typepad](#) and there are many others. Installable open source blog servers include [Wordpress](#) (PHP), Roller (Java). [Movable Type](#) is the dominant proprietary product, however there are many others including [Community Server](#) (.Net).

An important consideration for corporate and internal blogging is having a company policy that gives appropriate structure to employee's blogging behaviour. IBM bloggers wrote a policy and set of core principles using an internal wiki. This has advice such as *Know IBM's Business Conduct Guidelines, Be who you are, Speak in the first person, Use a disclaimer, Respect copyright and fair use laws, Protect confidential and proprietary information, Don't pick fights, Try to add value.* Sun has a set of guidelines including *It's a Two-Way Street, Don't Tell Secrets, Be Interesting, Write What You Know, Quality Matters.*

Another important consideration is policy regarding comments. Comments can be anonymous or authenticated, and moderated or unmoderated. Each permutation has different consequences. Authenticated comments mean users have to have an account on the system, and are identified as the originator of the comment. Moderated means that the owner of the blog approves comments before they appear publicly. Anonymous and unmoderated comments pose few barriers to contribution, but have a higher risk of being of an inappropriate or inflammatory nature. The combination of authenticated and unmoderated comments are often used to create self managing behaviour. There are cases where the preservation of anonymity can be positive, for example the CEO of the NZ Fire Service has an internal blog in which moderated comments are received and published anonymously to encourage free exchange of feedback and constructive criticism.

A new form of blogging that has recently emerged is ‘microblogging’. This is a post which is restricted to 140 characters and is typically used to tell others what you are currently doing or thinking about. People use it to share updates on their daily activities, breaking news items, and links to full blog posts. The microblogging phenomenon has been popularised by [Twitter](#), and is also used in [Facebook](#)’s status update.

### 3.2 Wikis

The wiki exists as a tool for collaborative content creation. Similar to a blog, it allows writing of content directly to a public space on the web or intranet. However, unlike a blog, rather than just commenting on the content the audience can modify and contribute directly to the content. By its very nature a wiki sees the author and the audience participant as the same thing, and it is the exception rather than the rule that an audience member will only have read access to the wiki. Wikis have been a dominant web based collaboration tool for a number of years, and have seen service in many environments and have been customised for many tasks. In recent times they have been the tool of choice for online reference stores, the best example being [Wikipedia](#), the ubiquitous online encyclopaedia. Its use within the corporate intranet environment as a KM tool is covered well by Dick Stenmark’s research paper [‘Knowledge Sharing on a Corporate Intranet: Effects of Re-Instating Web Authoring Capability’](#).

Wikis allow anyone to edit pages, but keep a complete track of version history. As such it is easy to compare differences across versions and to roll back changes. While it seems counterintuitive, this opening up of editing capability to everyone typically allows the best possible quality content to emerge. Where people spot something they believe to be incorrect they can fix it immediately, rather than going through a lengthy editorial process.

Generally wikis are seen as a fast and easy to use collaborative tool, however new users can find them frustrating due to wiki formatting. Content displayed in a wiki is formatted using a type of markup; users are expected to mark up portions of text using patterns of punctuation marks inline with content. The purpose of this markup is supposed to make wikis more accessible to people not familiar with HTML. The popular solution is training and good documentation, however some more recent wiki implementations provide editor interfaces that emulate word processors.

As with any mature technology there are plenty of options available from both the open source and proprietary communities. Most provide a try before you install/buy approach where the user can test the wikis capabilities on the provider’s site, which can be useful for weighing up the various editing interfaces.

[MediaWiki](#) is perhaps the widely used wiki software; it powers Wikipedia and is considered mature and easy to use. It implements a fairly primitive word processor style interface to help novices come to terms with the syntax, injecting the markup into the content being edited. It is relatively easy to install and configure, and although there is no formal support there is a sizeable community and plenty of documentation.

One of MediaWiki's biggest open source competitors in the enterprise space is [TWiki](#). Billed as an enterprise level wiki, TWiki differentiates itself by having a large selection of plugins that can be used to seriously customise the tool. The danger with plugins is that they tend to break very easily due to changes in the core wiki codebase or in some dependant technology. The main reason for this is that plugins tend to be third party contributions that are not maintained with the core codebase. Like MediaWiki there appears to be no formal support (which is odd for a tool pitched at the enterprise user), however there is an equally strong community and equally plentiful documentation. TWiki feels like it is geared toward users with development experience rather than everyday office staff, and seems to lack nice-to-have features like a word processor interface within its core package.

Of the licensed wikis on offer the dominant players are [Confluence](#) and [Socialtext](#). Both are well presented and feature rich wikis, it support exporting content from the wiki to offline formats like pdf (portable document format) and implement word processor style interfaces that hide wiki formatting and use the typical hot keys associated with Word and OpenOffice. They both offer formal support, and have seen adoption by many enterprise customers.

[Wikimatrix](#) provides a very comprehensive list of wiki software and provides feature for feature comparison tables.

### 3.3 Tags

Tags are a widely deployed information categorisation method in Web 2.0 technologies including blogs, photo sharing and social bookmarking applications. Tags are essentially keywords, however they differ from traditional controlled vocabulary keyword lists. Tags are user defined, and as such are not 'tidy' like traditional taxonomies. The structure they create is often referred to as a 'folksonomies'. Tags enable faceted classification systems to develop, where one item can be in many categories. This 'messiness' is seen as beneficial in that a useful and navigable order emerges bottom up, the structure can adapt more quickly to change, and can incorporate more fine grained distinctions. For example, people that tag things as 'movies' might actually have a different point of view from people who tag things as 'cinema'.

When an information item (such as a blog post) is tagged, that tag generally links to all other articles tagged in the same way (either by that blogger or a wider group). Tags are often presented in tag clouds where the size of each tag represents the number of items it represents. Examples of tag clouds can be seen at <http://www.technorati.com/tag/>, <http://www.flickr.com/photos/tags/>, <http://del.icio.us/tag/>, and closer to home on <http://www.nzhistory.net.nz/> and <http://www.nzlive.com/>. Tags are scalable ontologies in that they give a user complete control over the way they sort their own information, but when merged with other users' tags they automatically create an aggregated categorisation scheme.

Tags can be expert created, or user created, or both. It is typical to open up tagging to the widest possible audience, however it is also possible to differentiate between the sets of tags created by different groups of people. For example distinctions could be made between tags created by experts (internal or external) and regular users, or between different groups within an agency, or between different types of users (e.g. researchers, genealogists, students).

### 3.4 Social Bookmarking

Social bookmarking is a popular new way to store, classify, share and search links to web sites. Rather than storing links in one's browser, these services allow storage of bookmarks on a server. They use tagging to create easily navigable user defined classification of bookmarks. Users can see what other people tagged a site, and what other sites people bookmarked for particular tags.

Extensions for some modern web browsers like Firefox 2.0 integrate social bookmarking services directly into the browser's bookmarking interface.

Popular social bookmarking sites include del.icio.us, Furl, and Connotea. A good analysis of the social bookmarking phenomenon can be found at <http://www.dlib.org/dlib/april05/hammond/04hammond.html>.

### 3.5 Social Networking Services

Social networking services are web sites which allow people to make explicit connections to people they know, and easily share information only with those people. Users often share their likes and dislikes, their photos, blog posts, short descriptions of what they're currently doing or feeling, and use the systems to publicly or privately send messages to each other.

Popular social networking sites include Facebook, MySpace, Bebo and LinkedIn. These services are starting to be used in the commercial world, in particular LinkedIn (which has a professional focus) and Facebook. Teams in businesses are using Facebook accounts to keep up to date with each other. LinkedIn is commonly used to ask questions of one's network, or seek referrals to sales or employment prospects through 'trust tagging' (i.e. discovering someone you want to meet is known by one of your acquaintances).

## 4 Use at Archives NZ

There are many different ways Archives NZ could use Web 2.0 technologies and approaches. These include internal use for enhanced collaboration and knowledge management, and external use to both harness community contribution to archives, and to support other government agencies in sharing knowledge about good record keeping practices.

### 4.1 External Use

#### 4.1.1 By the Public

Archives NZ holds content that is regularly accessed by external users for many purposes. This content is expected to grow significantly as digitisation programmes and the storage of digital born content progress. The Ministry of Culture and Heritage are currently experimenting with a number of approaches that enable users to both describe and contribute content to NZ History, NZ Live, and Te Ara. Archives could take a similar approach with content stored in Archway. This could harness a huge voluntary labour force to enhance the discoverability and usefulness of Archives' content. This could include:

- Allowing users to edit metadata for records (perhaps in a moderated and change controlled way using wiki style technology)
- Allowing users to tag records to create user defined 'folksonomies'
- Letting users 'geocode' public records and contributed items (i.e. provide geospatial references so records can be visualised using mapping software such as Google Maps and Google Earth)
- Grouping, mining and presenting user contributed metadata to create aggregated views that may assist other users to locate records (e.g. in tag clouds)
- Letting users save searches, annotate their saved searches and make them available to other users
- Aggregating usage patterns (e.g. most viewed records, most common search terms) and displaying these on the site
- Using collaborative filtering methods to display 'people who viewed this record also viewed these records'
- Allowing users to attach other content to records, either by uploading or linking to external content. For example users annotate records (either using text or audio) and add images, or could link records to external web sites, blog posts, or photos (or associated tags and other metadata) on photo sharing sites such as Flickr
- Letting users contribute their own records
- Providing functionality so users can create adhoc online collaborative communities around particular topics of interest. This could be on the Archives NZ site, but could leverage web services provided by social networking sites such as Facebook and photo sharing sites such as Flickr.
- Enabling Archives NZ staff to participate in external user communities to provide expert advice and support

### 4.1.2 By Government Agencies

Archives NZ has the role of supporting agencies in meeting the requirements of the Public Records Act. Web 2.0 approaches could be used to support this by:

- Providing online forums and communities for records managers in public sector agencies
- Using wikis to establish emerging/good/best practice guidance, tips, and FAQs for public sector records management
- Using blogs to disseminate news, ideas and emerging thinking on records management practices, and allow users to provide comments and feedback
- Host 'guest bloggers' including external records management experts, and experienced/innovative public sector practitioners

### 4.2 Internal Use

There are many uses Archives NZ could make of Web 2.0 technologies to enhance internal collaboration and knowledge management. Moreover, Archives NZ should consider pioneering and experimenting with these approaches internally before trying them externally. Uses could include:

- Providing wikis for aspects such as policy & procedure documentation, FAQs, how-to manuals, project documentation, and collaboratively authored documents
- Using blogs for CEO and senior management newsletters, and enabling anonymous or authenticated comments
- Allowing staff to blog internally about professional topics of interest to them. Staff could also be permitted to blog externally.
- Providing for tagging of internal content (on the Intranet, document management system, or of archival content)
- Using microblogging to help staff keep up to date on each others' activities
- Implementing social bookmarking so staff can see each other's bookmarks
- Providing an Intranet home page that staff can personalise by adding feeds from internal and external sources (e.g. their favourite blogs)
- Enabling users (if they grant permission) to let other staff see their personalised set of feeds
- Allow staff to create Facebook style profile pages on the Intranet

### 4.3 Risks

There are a number of risks that Archives NZ may face in adopting Web 2.0 technologies and approaches. These are as follows.

Risk	Mitigation
Archives NZ staff are resistant to adopting these approaches	Prototype technologies rather than doing all at once Establish champions and early adopters
Archives NZ culture and incentive/reward structures do not support adoption of Web 2.0 approaches	Clarify level of senior management buy-in to Web 2.0 philosophy and their support for associated initiatives and tools
Some approaches fail and staff become disheartened or cynical	Study what has worked elsewhere Communicate to staff that not everything we try will work Try lots of things and keep what works
Content generated by the public is inappropriate or offensive	Develop and promote appropriate policies Experiment with moderation (by staff or by valued/trusted members of the community of public users)

### 4.4 Culture Change

In order to use Web 2.0 technologies externally it may be necessary to affect culture change internally. It is not uncommon for people who have spent much of their professional careers using traditional methods of taxonomy, records keeping and information management, to be cautious of, or resistant to approaches which are inherently messier and less ordered.

Embracing Web 2.0 methods, and harnessing the power of community contribution means being willing to give up some control. The Archives NZ culture must support this if these kinds of initiatives are to succeed. Peter Van Garderen in a post<sup>1</sup> on his archivemati.ca blog says:

*“archival institutions are going to have to accept the rise of grassroots archivists. Not as barbarians at the city gates but as value-adding partners that share the goal of preserving historical memories and experiences. In his excellent webcast presentation, [Are the Archives Doomed?](#), Rick Prelinger discusses the emergence of what he calls ‘archives groupies’ and the wonderful, often unexpected results that occur when users are invited to participate in the organization and use of archival collections.”*

In exploring the use of Web 2.0 approaches it is very difficult to predict what will work, and what won't. The best method in this context is to try many things and keep those that are successful. This requires a culture that is tolerant of failure. It must be acceptable for initiatives not to work, as long as people learn from them and adapt as a result.

For Archives NZ to be a credible advocate of good information management to other government agencies it must practice what it preaches. Ideally Archives NZ should have best practice processes in this area, and be an innovator/early adopter, trying out new approaches. This requires a culture that is able to experiment, learn quickly, and embed successful approaches into stable best practice.

<sup>1</sup> <http://archivemati.ca/2006/05/08/web-20-and-archival-institutions/>

In experimenting with Web 2.0 approaches internally it will also be important for Archives NZ to not assume that because something didn't work internally, it won't work externally. Internal experimentation should be done to understand new technologies and approaches, rather than to prove them before deploying. An experimental approach will also be necessary externally, but it should be done with the knowledge gained through internal use.

## 5 References and Further Reading

Thanks to Tim Barrable, CIO, Ministry for Culture and Heritage for his input into this report.

### Further references are:

Corporate Blogs: Weapons of Crass Discussion?

[http://www.intranetjournal.com/articles/200605/pij\\_05\\_24\\_06a.html](http://www.intranetjournal.com/articles/200605/pij_05_24_06a.html)

IBM blogging policy and guidelines

[http://www-03.ibm.com/developerworks/blogs/page/jasnell?entry=blogging\\_ibm](http://www-03.ibm.com/developerworks/blogs/page/jasnell?entry=blogging_ibm)

Sun Policy on Public Discourse

<http://www.tbray.org/ongoing/When/200x/2004/05/02/Policy>

Knowledge Sharing on a Corporate Intranet: Effects of Re-Instating Web Authoring Capability

<http://csrc.lse.ac.uk/asp/aspecis/20050001.pdf>

Social Bookmarking Tools

<http://www.dlib.org/dlib/april05/hammond/04hammond.html>

Are Archives Doomed?

<http://archivemati.ca/2006/03/10/are-archives-doomed/>

Web 2.0 and Archival Institutions

<http://archivemati.ca/2006/05/08/web-20-and-archival-institutions/>

### Wikipedia definitions:

<http://en.wikipedia.org/wiki/Blog>

<http://en.wikipedia.org/wiki/Wiki>

<http://en.wikipedia.org/wiki/Tags>

[http://en.wikipedia.org/wiki/Social\\_network\\_services](http://en.wikipedia.org/wiki/Social_network_services)

### Blogging Software:

Blogger – <http://www.blogger.com>

Wordpress - <http://www.wordpress.com>

Moveable Type – <http://www.moveabletype.org/>

Typepad - <http://www.typepad.com/>

Community Server - <http://communityserver.org/>

DotNetNuke – <http://www.dotnetnuke.com>

### Wiki Tools:

Mediawiki - <http://www.mediawiki.org/wiki/MediaWiki>

Twiki - <http://twiki.org/>

Confluence - <http://www.atlassian.com/software/confluence/>

Socialtext - <http://www.socialtext.com>

WikiMatrix - <http://www.wikimatrix.org/>