

## CHAPTER 1

# The big picture: Web 2.0 and current trends in IT

## **Q**uestions addressed in this chapter

- What defines Web 2.0?
- How does it differ from Web 1.0 and what do they have in common?
- What are some of the other trends in IT and information creation that are influencing, or influenced by, the rise of Web 2.0?

### **What is Web 2.0?**

---

You don't need to be an IT professional to have noticed the rapid rise of a new breed of web applications and services over the past couple of years. Names such as YouTube, Facebook and Wikipedia have seemingly appeared from nowhere to become part of the cultural mainstream almost overnight. And these represent just the tip of a very big iceberg, one which is commonly known as Web 2.0.

It must be acknowledged at the outset that this is not a term or even

a concept accepted by all and is indeed disputed by such luminaries as the man credited with the invention of the world wide web itself.<sup>1</sup> But equally, there exists a growing body of evidence and band of supporters who perceive a significant enough change, both in the technology and in the way in which it is being utilized, to support the assertion that what we are now witnessing does indeed represent a new, second iteration of the world wide web.

That said, the scepticism of Tim Berners-Lee and others should at least serve to remind us that any definition of Web 2.0 is bound to be imperfect, open to challenge and likely to change over time – so too the type and range of services and applications that may be said to fall under its umbrella.

It is possible to identify at least seven distinct types of Web 2.0 application (van Harmelen, 2007) and in many respects it is through such a list that we can most usefully define what we currently mean by Web 2.0.

According to van Harmelen these seven types of Web 2.0 software include:

- 1 Blogs
- 2 Wikis
- 3 Social bookmarking
- 4 Media-sharing services
- 5 Social networking systems
- 6 Collaborative editing tools
- 7 Syndication and notification technologies.

But while the above list tells us something of the type of web service that can be described as ‘Web 2.0’, it does little to clarify what it is that actually differentiates them from their Web 1.0 predecessors. To do this we need to focus our attention away from the underlying technology

and towards the respective role of the user in both the Web 1.0 and Web 2.0 worlds.

Since its invention in the early 1990s the user experience of the world wide web has been a largely passive one. Within most organizations web content has been produced by a relatively small and limited clique comprising those with the technical ability to produce web pages and the authority within organizations to publish ‘approved’ output. In this respect, though the delivery mechanism was very different, the basic model underpinning Web 1.0 closely mirrored that of traditional publishing, with content being provided by a limited number of sources and the consumer having little choice but to either read, or ignore, the information being provided. There was comparatively little potential for active or innovative reuse of the material and what feedback mechanisms existed were limited, stilted and ‘artificial’.

Web 2.0 has turned this model on its head; the floodgates are well and truly open and we are all potential content creators now. Sites such as YouTube, Flickr and Wikipedia not only encourage the individual user to contribute content, they are totally reliant upon it. Without the tens or hundreds of thousands of individuals choosing to upload their favourite video clips or images, there would be no content. But it doesn’t end there. For every user who contributes content there might be another dozen who comment on it, rank it, link it to related material, take it and embed it into their blog, or reuse it for another purpose. Web 2.0 is a world without walls where content is king and the potential for its use and reuse limited only by the imagination of the individual user.

Rather more surprisingly, these services are nearly always provided free of charge – at least for the basic level service. This is something we are rapidly taking for granted but which is a radical departure from the accepted rules of business. This has undoubtedly been a decisive factor in the explosion in growth that such services have witnessed in the past couple of years. It has helped form the massive

user base that is an essential prerequisite of all Web 2.0 applications and services, reliant as they are not only on the presence but on the active participation of massive numbers of individual users, sometimes referred to in this context as ‘the crowd’.

It is for this reason that the records management community cannot afford to turn a blind eye to the implications of Web 2.0. The web is no longer a passive publishing vehicle, something which we can take advantage of for our own purposes but largely dismiss as an area of professional concern. Just as we have long acknowledged and agonized over the fact that our users conduct business transactions via e-mail and therefore create records in e-mail (despite its unsuitability as a technology for the purpose),<sup>2</sup> so the same will be true with Web 2.0. If the defining characteristic of Web 2.0 is the creation, sharing and repurposing of information by individual users we have to acknowledge that this surely places this technology and its use firmly and squarely in the back yard of the records management profession.

### **Similarities and differences compared to Web 1.0**

---

But if it is this notion of the web as a platform, rather than just a publishing vehicle, that most differentiates it from what has gone before, this does not mean that Web 2.0 has nothing in common with its predecessor. Many of its characteristics have actually been an integral part of the world wide web since its inception over a decade ago and are likely to continue to be equally pertinent in the Web 2.0 world. Upon first glance some of these qualities appear so obvious and self-evident as to be unworthy of comment but, as we shall see, though often previously taken for granted, their significance to the management of information should not be underestimated – a significance which will, if anything, increase with the rise of Web 2.0.

Firstly, the vast majority of the content and services available online

via the web are generated and provided by external agents over which your organization has no control. When the web was simply a source of reference material it could be argued that this was little different from the pre-online era, where staff may have relied just as heavily on printed journals, magazines, text books and broadcast media – all of which too originated from external sources. But now that these services are also providing the *applications* being used by staff, and *storing* the content they create, the implications of this may be far more significant.

One of the consequences of the remote nature of the web is that it requires little by way of client technology to function. A PC or laptop, a modem, a browser and some anti-virus software are the minimum required. This is a very different model to the application-heavy burden placed on the user's PC by the standard office software suite, where operating systems and software applications have to be physically installed and located on the local hardware. This is, of course, a feature closely related to the external nature of the web and, as we shall see, it is responsible for further weakening the ties of dependency that currently bind users to the internal IT infrastructure of their organization.

A final characteristic shared by both the Web 1.0 and Web 2.0 worlds is their transience. Some researchers claim the average life expectancy of a web page to be less than two months, while the bursting of the dot-com bubble in the late 1990s taught many investors the hard way that little in the online world possesses the kind of stability and longevity that breeds long-term trust. Even in the short period of time that Web 2.0 has been making the news it has been possible to see the fortunes of individual services wax and wane. A year ago MySpace was seen as the great online phenomenon; now that crown is (at the time of writing at least) being worn by Facebook. But in this fickle and fast-moving world how long will it be before this too is eclipsed by some new and currently unknown social software wunderkind? We have already identified that Web 2.0 services rely on the presence and participation of the crowd to function. Whereas in

the Web 1.0 world a drop off in visitor numbers might have caused a slow steady decline in advertising revenues that might eventually have threatened the long-term viability of traditional websites, the desertion of the crowd is likely to have a much more sudden, profound and irrevocable impact on the prospects of Web 2.0 services.

This will naturally be a key cause of concern for the records manager. For as a profession we are used to taking the long view, not only looking at the value of information for today, but at its potential relevance in the future. We are rightly careful to ensure that we take whatever measures are needed to ensure the survival and accessibility of records in the future, be that by the creation of back-up copies, by format migration, or by ensuring the suitability of the physical storage environment. Such measures help maximize the chance of survival of the records we believe we will need into the future, and are largely possible because the items in question are within our possession and held within our own hands (be they physical or virtual). What impact will the reliance of our users – and potentially our organizations – on external service providers have on this ability, and what additional risks may it create in an uncertain technical and commercial world?

### **IT trends: blurring the boundaries**

---

‘Thanks. Interesting.

But shouldn’t you put these up on YouTube, where they can be commented on, embedded and linked to, like everyone else does?’

This was one immediate e-mail response received by an organization after publicly announcing that it had just made two short research and development videos freely available on its website; though brief it encapsulates an interesting new by-product of the Web 2.0 world.

Sites such as YouTube and Flickr have now attained such market

dominance that it is rapidly becoming inconceivable that any organization should bother to reinvent the wheel and develop its own facility for storing and sharing its own content. It is clear from the above quote that the author of the e-mail considers YouTube the natural home for online video clips. The chances are that he (or she) uses it extensively for both domestic and work purposes. He is likely to be an active account user who is familiar with the service offered and, more importantly, knows that he can make use of other Web 2.0 applications to repurpose the content and make maximum use of it in a variety of different contexts.

Evidence is emerging that this is not a phenomenon restricted to media-sharing services such as YouTube or Flickr. A recent report commissioned into the use of Web 2.0 within the UK Higher Education sector uncovered examples of academics who are eschewing the formal systems provided by their institutions, such as the Virtual Learning Environment, and are choosing instead to use Facebook as their preferred means of keeping in contact with their students (Anderson, 2007).

Such examples not only help demonstrate that Web 2.0 applications are already beginning to have a very real impact on some sectors, but also how they are beginning to blur the boundaries between domestic and business use of IT. It is achieving this by becoming users' first and preferred service in both spheres of their life. Of course this has been true for many years in terms of people's common use of Microsoft Office applications or Google both at home or in the office, but that was very different. In these examples the applications may have represented common tools, but were tools largely being used for very different and distinct purposes. By using them to write a budget report at work and a letter to their child's school at home the traditional boundaries between office and domestic life remained largely intact. Plus, of course, though it may have been the same product being used in both examples, the chances are that it will have been two physically different instances of that product in use.

The crucial difference with Web 2.0 services such as YouTube, Flickr, Facebook and the like is that they are content storage repositories as well. They no longer just represent the tools, but also the filing cabinet: filing cabinets that now contain side by side the combined outputs from both our domestic and work life. What is more, because of the core concepts of data reuse and manipulation underpinning Web 2.0, users are now likely to make use of content created in one part of their life as part of their role in the other without a second thought (for example using a photograph taken during a weekend family outing and stored on Flickr to illustrate a point in a presentation they create the next day for a team meeting at work).

### **IT trends: the exponential age**

---

‘In a world with infinite storage, bandwidth, and CPU power, here’s what we could do with consumer products . . . Store 100% of User Data. With infinite storage, we can house all user files, including: e-mails, web history, pictures, bookmarks, etc and make it accessible from anywhere.’<sup>3</sup> Volume is now everything. The above quote from a leaked internal Google presentation demonstrates how, for Google, the seemingly boundless growth in storage capacity has the potential to revolutionize the world.

Most of us will be familiar with Moore’s Law (the ‘law’ first articulated by Dr Graham Moore in 1965, revised in 1975, which states that the number of transistors on a chip will double every 24 months),<sup>4</sup> which has proved a remarkably accurate prediction of the growth in storage capacity over the past 30 years. The even more powerful and unbeatable laws of physics must surely dictate that one day Moore’s Law must fail – that we shall reach a point at which it is physically impossible to cram ever more technology into even less space, or that we shall run out of the energy required to power the vast server farms

required by companies such as Google. But there are few signs that we will reach either point any time soon. In fact, every time the chip industry seems to run up against a seemingly insurmountable obstacle to keeping pace with Moore's prediction, a new technical breakthrough occurs to keep it on track: for example with the recent replacement of the silicon dioxide traditionally used to make the gate dielectrics within transistors with the metal hafnium – a change described by Graham Moore himself as 'the biggest change in transistor technology since the late 1960s'.<sup>5</sup>

Evidence of the impact of continual storage growth is all around us and shaping our daily lives. Take the popular Yahoo! e-mail service for example. When it launched in 1997 it offered its users 4Mb storage for their e-mail accounts; by 2004 that figure had soared to 100Mb. In May 2007 the company took the next logical step and now offers all of its users unlimited e-mail storage. The vision of the future described by John Kremer, Vice-President of Yahoo! Mail, when announcing this development has clear parallels with Google's view of where we are heading: 'We hope we're setting a precedent for the future. Someday, can you imagine a hard drive that you can never fill? Never having to empty your photo card on your camera to get space back? Enough storage to fit the world's music, and then some, on your iPod? Sounds like a future without limits.'<sup>6</sup>

The participation of the crowd, mass media storage and the ability to continually repurpose content – all inherent characteristics of Web 2.0 – would not be possible without access to the levels of storage capacity we now take for granted.

These technical developments are also having an impact far beyond the world of IT and records management, and represent one of the most powerful forces shaping popular culture and user behaviour in today's world. Ten years ago a fortnight family holiday meant taking a couple of 24-exposure camera films and doing your best to spread them out over the two weeks. Now we can take that amount of pictures or more

every hour and, should we wish to, choose to keep every image.

Then there is the web itself, providing the user with free and instant access to several billion pages of information. People may occasionally complain that they can't find what they are looking for, but seldom call for it to be reduced in size. Thanks to Google and the like we now have a new generation emerging who would far rather *search for* than *manage* information, and who value quantity over quality. Indeed this trend is even beginning to influence the information management profession itself, as the following quote from *Information World Review* makes clear: 'Information is inherently disorganised and not uniformly stored, and it is better to provide good search than to provide good organisation of the information.'<sup>7</sup>

If we were to accept this train of thought, where would this leave the records management profession, a profession based on an underlying assumption of centralized control? Indeed, given that most of the central tenets of records management practice were defined in the pre-electronic age when the volume of records being created was a fraction of that created today, it begs the question: is records management scalable enough?

It is not uncommon to hear information professionals rather smugly dismiss Google and its inaccuracy, pointing to the fact that every search brings back several million web pages in its results, and comparing this unfavourably to their tightly defined, metadata-controlled search engine which only returns a handful of relevant hits. But this is to miss the point entirely. The user couldn't care less about the six million other results that Google has returned, because almost without fail the information they are looking for appears at, or very near, the top of the list. The number of times the user needs to refer to the second or subsequent pages of results is, in my experience, rare: and this is through simple basic searching. What is more, it has required no effort on behalf of the content creator, or the end user. No need to follow metadata schemas, consider classification schemes, or add additional keywords.

All the while Google focused on being a web search engine the genie was still just about in its bottle. Users may have been happy to rely on Google to retrieve and rank four million hits when it came to the internet, but were still – within reason – willing to think of their internal documents in a different light and to continue to file things according to their organizational file plan and add metadata where necessary.

A host of new desktop search engines such as Copernic<sup>8</sup> and Google desktop<sup>9</sup> are rapidly changing this and bringing the same logic of cheap storage and easy retrieval to the contents of our PCs and laptops. Users are now able to see just how easy it is to retrieve their files, regardless of the folder they have put them in, or what they have been called, and with absolutely no decision making or manual input required on their behalf.

If more proof where ever needed that volume and retrieval technologies are the undisputed drivers behind IT, take a look at the companies listed on the Nasdaq-100 Technology Sector Index.<sup>10</sup> The majority of the IT behemoths listed either relate to storage (Intel, SanDisk, Oracle, Sun, etc.) or, to a lesser extent (thanks to Google's dominance one suspects), search. Far less prominent are those companies dealing with the classification, management and disposal of information. Put simply: storage sells; disposal doesn't!

As the American blogger Karl Fisch succinctly puts it: 'We are living in exponential times.'<sup>11</sup> His excellent presentation *Did you know?* contains a barrage of fascinating statistics to ram home this point, one of the most telling of all being 'The amount of new technical information is doubling every two years . . . it is predicted to double every 72 hours by 2010'<sup>12</sup> – a sobering statistic which reminds us that what we are seeing today, though profound in itself, is but the tip of the iceberg that lies ahead.

## References

---

- Anderson, P. (2007) *What is Web 2.0? Ideas, technologies and implications for education*, [www.jisc.ac.uk/media/documents/techwatch/tsw0701bword.doc](http://www.jisc.ac.uk/media/documents/techwatch/tsw0701bword.doc) [accessed 3 March 2007].
- van Harmelen, M. (2007) *Briefing Paper on Web 2.0 Technologies for Content Sharing: Web 2.0, an introduction*, <http://franklinconsulting.co.uk/LinkedDocuments/Introduction%20to%20Web%202.doc> [accessed 5 October 2007].

## Notes

---

- 1 A transcript of a podcast interview of Tim Berners-Lee, inventor of the web, [www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.txt](http://www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.txt) [accessed 5 October 2007].
- 2 Due largely to the range of formal and informal purposes that e-mail is commonly used for, the nature of most e-mail applications (which makes the acts of creation and replication far easier than management), and the common misconception that this has led to that e-mail is an ephemeral and untraceable medium.
- 3 <http://blog.outer-court.com/archive/2006-07-11-n52.html> [accessed 16 July 2007].
- 4 [http://en.wikipedia.org/wiki/Moore%27s\\_law](http://en.wikipedia.org/wiki/Moore%27s_law) [accessed 2 March 2008].
- 5 <http://news.bbc.co.uk/1/hi/technology/7002267.stm> [accessed 19 September 2007].
- 6 <http://ymailuk.com/blog1/2007/03/28/yahoo-mail-goes-to-infinity-and-beyond/> [accessed 20 September 2007].
- 7 Beyond Search, *Information World Review*, **229** (November 2006), 2–21.
- 8 [www.copernic.com/](http://www.copernic.com/) [accessed 3 March 2008].

- 
- 9 <http://desktop.google.com/> [accessed 3 March 2008].
  - 10 [www.nasdaq.com/services/indexes/ViewIndexes/Nasdaq\\_tech.aspx?symbol=NDXT](http://www.nasdaq.com/services/indexes/ViewIndexes/Nasdaq_tech.aspx?symbol=NDXT) [accessed 2 February 2008].
  - 11 <http://thefischbowl.blogspot.com/> [accessed 2 February 2008].
  - 12 <http://thefischbowl.blogspot.com/2006/08/did-you-know.html> [accessed 2 February 2008].