Web 2.0: Hypertext by Any Other Name?

David E. Millard and Martin Ross
Electronics and Computer Science
University of Southampton
Southampton, UK
+44 (0) 23 8059 3255
{dem, mar302}@ecs.soton.ac.uk

ABSTRACT
Web 2.0 is the popular name of a new generation of Web applications, sites and companies that emphasis openness, community and interaction. Examples include technologies such as Blogs and Wikis, and sites such as Flickr. In this paper we compare these next generation tools to the aspirations of the early Hypertext pioneers to see if their aims have finally been realized.

Categories and Subject Descriptors
H.5.4 [**Hypertext/Hypermedia**]: Theory

General Terms
Design, Theory

Keywords
Web 2.0, Hypertext Pioneers, Hypertext Functionality

1. INTRODUCTION
The World Wide Web was originally one of a number of hypertext systems proposed and promoted in the early 1990's [13, 1, 11]. Through its philosophy of simple architecture and open standards, and benefiting from the network effects of its global design, it has since grown into a vastly distributed and diverse information and application platform.

The most recent generation of Web applications and Web sites have been considered by some to be fundamentally different from the ones found on the early Web, these have been grouped together under the term Web 2.0, and while the name is arguably misleading (implying a designed version and a discrete evolution) the concepts beneath it provide a valuable insight into the way in which the Web has evolved.

O'Reilly characterises Web 2.0 applications, websites and companies by example and discussion [20], comparing Web 1.0 efforts such as content management systems, directories and screen scrapping with Web 2.0 alternatives such as Wikis, Tagging and Web Services. The Web 2.0 concept is probably still too intangible for a solid classification, however it can be said that the Web 2.0 approach emphasises interaction, community and openness.

In 1945 Vannevar Bush published his seminal paper, "As We May Think" [6], reflecting on how technologies could help solve the problems of post war society. He is now considered to be the inventor of hypertext, and ever since his original work others have been refining his approach, both in terms of what hypertext should offer, how it is interpreted by readers, and also in concrete systems and studies.

Walker has noted that the new generation of Web applications has created a new form of feral hypertext, unrestrained by systems or ownership [22]. In this paper we analyse the field of hypertext research in order to draw out the aspirations of its pioneers and their subsequent refinements by the community, we then compare these aspirations with a number of Web 2.0 systems in order to draw conclusions about how well those earlier ideas have been realised in the modern Web.

2. ASPIRATIONS
In 1987, Halasz took the aspirations of the earlier pioneers and, based on the systems around at the time, described seven issues that had to be considered and resolved in order to progress towards the systems envisioned; these included composite structures, versioning, collaboration and search [12].

In the 1990s there was an effort to create standards for hypermedia systems both in the Web community but also in the Open Hypermedia community, represented by the Open Hypermedia Systems Working Group (OHSWG) [8]. Using standardized protocols and an open architecture, Open Hypermedia systems separate links from content allowing the provision of many functionalities envisioned by early researchers. The same principle can be seen in many web systems today, where links are generated when required (usually from a database) rather than embedded into static pages.

Combining the original aspirations with the issues Halasz describes, as well as taking into account new issues raised in the literature since, allows us to identify a core set of aspirations that we can use to compare Web 2.0 systems. The following aspirations are not meant to be original, or exhaustive, but we believe capture many of the important technical challenges that have been identified by the hypertext research community.

2.1 Search
A key aspiration is that hypertext systems should allow people to search for both content and structure. Content may be multimedia, which would require new matching techniques [4]. Examples of structures in the hypergraph include patterns such as cycles or mirrorworlds [3]. Finally context may be an issue, does the user mean 'pie' in the sense of food, or statistical charts [10]? This gives is three aspirations:

Content Search – Content based retrieval, searching for a particular word pattern or multimedia object

Context Search – Typically using meta-data to distinguish between categories of searchable entities

Structural Search – searching for a particular pattern within a hypertext graph

2.2 Structure and Content

Hypertext can include structures beyond a point-to-point link. Links can be n-ary and they can be typed, allowing semantic networks to be modeled – crucial for domains such as argumentative hypertext [7]. Other structures are useful, such as guided tours, and nodes may be comprised of many different resources gathered together in a single composite (analogous to transclusions in Xanadu [17]). Dynamic structures such as user trails are useful to allow people to effortlessly add to the hypertext, and implicitly value its content [21]. From this we can extract four aspirations:

Typed n-ary links – multi-headed, typed links, so an anchor can be semantically associated with more than one destination.

Composition – combining several entities into one collective entity that can be referenced as if it were atomic.

Extended Navigational Structures – non-link structures for advanced browsing, such as guided tours or categorizations.

Trails – Keeping a track of a user's history to enable recommendations or simple tour authoring

2.3 Dynamic/Adaptive

Hypertext systems should not only deal with static content or structure, dynamism allows hypertext systems to be coupled with live information and affords adaptation and personalization [9]. We can distinguish four aspirations:

Dynamic Content – dynamic creation of entities and content within them, perhaps drawing on live data (such as a DB)

Dynamic Structures – dynamic creation of structures, from simple links, to collections, tours or other structures.

Computation over the network – performing calculations within the hypertext (e.g., triggered by link following).

Personalization – Facility to allow a user to personalize their view of globally available documents and links.

2.4 Versioning

Versioning has long been important in hypertext and information systems [18], in hypertext we can distinguish between versioning the content of nodes and the hypertext itself [14], giving us two further aspirations:

Entity Versioning – Storing a retrievable history of a node, so changes can be reversed, and branches managed

Network Versioning – Taking into account links to other nodes and effectively versioning entire (sub)networks

2.5 Open Authoring

The key to many hypertext systems is the lack of distinction between authors and readers, allowing all users to participate in a hypertext and blurring traditional literary and journalistic roles. Augmenting a live hypertext is often termed annotation and can be done either publicly or privately [15]. Live authoring leads to the possibility of collaboration and collaborative texts. We can extend this to the system itself by allowing users to extend the systems functionality [19]. We can thus factor five aspirations:

Private Annotation – Allowing a single user, or defined group, to augment a personally owned node with comments.

Public Annotation – Allowing any user of the system to add annotations to a public hypertext.

Global Collaboration – The ability for all users to productively work on the same documents independently

Restricted Collaboration - The ability for groups of people to productively work on the same documents independently

Extensibility – Ability for users to extend the functionality of the system, and for the system to grow / scale.

3. THE SYSTEMS

Web 2.0 is not a system, nor even a class of systems, so for our comparison we have to take exemplary technologies and web sites and compare them to the aspirations in order to get an indicative view.

We have also looked at two OHS inspired systems that use the web to deliver open hypermedia. We have done this in order to get some feel of the distance between what is practiced in academic systems and what is the reality of the next generation of web tools.

The first Web 2.0 system we have analyzed is Flickr[1], this is a community site where users can upload and share pictures and tag them with metadata to help other users find images of things that they are interested in – it is often associated with Folksonomies (ontologies that have evolved from community practice). Wordpress[2] is a free blogging system written in PHP, like many blogging systems it supports newsfeed standards such as RSS and ATOM.

We also looked at two Wiki systems. MediaWiki[3] is the PHP wiki engine used by Wikipedia, it allows media content to be uploaded for inclusion in Wiki pages and supports discussion and alternative content namespaces. TWiki[4] is another free Wiki system, written in Perl, it has a strong authentication and versioning engine, which can be used to record and control authoring.

Finally we considered Annotea, a W3C project to define metadata standards for user annotation of Web resources [15], the current client implementation is called Amaya, although there are other efforts to include it natively in browsers such as Firefox.

---

[1] Flickr: http://www.flickr.com

[2] Wordpress: http://www.wordpress.com

[3] MediaWiki: http://en.wikipedia.org/wiki/Media_Wiki

[4] TWiki: http://twiki.org/

	Web 2.0 Flickr	MediaWiki	TWiki	WordPress	Annotea	Academic / Research Xspect	OHS/WWW
Content Search	X	X	X	X	X	+	X
Context Search	X	+	+	X	X	+	X
Structural Search	+				X	+	+
Trails	X			X			
Composition	X	X	X	X	X	X	X
Dynamic Content	X	X	X	X	X	X	+
Dynamic Structure	+			X	X		+
Typed n-ary links						X	X
Other Navigational Structures	X	X	X	+	+	X	X
Computation over the Network	X	X	X	X		X	X
Entity Versioning		X	X				+
Network Versioning			X				
Private Annotation	X	X	X	X	X	X	+
Public Annotation	X	X	X	X	X	X	+
Global Collaboration	X	X	X	X	+	+	+
Restricted Collaboration	X	X	X	X	+	+	+
Personalisation	X	+	X	X	X	X	X
Extensibility	X	X	X	X	+	X	X

Table 1: Hypertext Aspirations mapped against Web systems (X represents full support, + represents partial support)

We have also looked at two research systems. The first is Xspect, this is an XLink based system that takes OHS links and translates them into XLink, it can then render these in browsers using CSS to convert the XLink structures to HTML and Javascript [2]. The second is an older OHS Web implementation in the form of the DHM/WWW system (based on the DHM Open Hypermedia System). DHM/WWW uses embedded Java applets to add Open Hypermedia Links to Web pages [11].

4. ANALYSIS

Table 1 shows the mapping between the aspirations and the systems. Many systems have partial support of the aspirations. We have judged support as partial if a system does not entirely fulfil an aspiration, but does implicitly address it (for example, Annotea, which partially supports collaboration by allowing users to read and respond to each others annotations). We make the following observations:

- Search is important. Companies such as Google have made search the de facto interface to the web, and its pre-eminence is reflected here. However, the emphasis on content means that structural search is less well supported.
- Dynamic content and composition are ubiquitous. There is far less support for dynamic structure. This reflects the way in which content is often assembled from live sources, but structure is authored and static. When dynamic structure does appear it is often the result of systems supporting the creation of virtual collections.
- There is no support for typed n-ary links except in the research systems. These are possible to implement on the Web in a number of ways (including XLink) and the lack of support probably indicates that they are not needed in the interface layer. This may be because of a tendency to follow a 'translate to hypertext' approach [16] that models semantics in the underlying system, not in the hypertext.
- Navigational structures are diverse. For example: topic collections and bookmarks in Annotea, sets of photos in Flickr, namespaces in MediaWiki, revision list in TWiki, and categories in Wordpress. The structures have evolved as and where they are needed, rather than from a grand scheme.
- There is very little support for versioning. Wikis use versioning (perhaps due to their public authoring approach), but other systems seem to regard their hypertexts as transient and do not bother to version.
- Trails are similar, and are supported where they are judged to be useful, but not when they are considered an overhead.
- All the systems allow private and public annotation, and support collaboration. This is all part of the Web 2.0 philosophy of trusting the wisdom of crowds, and fostering value through participation. This approach is directly in line with the aspirations of the early hypertext community.

- All the systems reflect personalization. However the adaptation models are simple, for example basic filtering, and do not reflect the user modelling and sophisticated reasoning of some research systems.
- Extensibility is universally supported. Just as Web 2.0 encourages participation through usage, it also encourages it through the extension of the system itself by the community.

5. CONCLUSIONS

In this paper we set out to show whether Web 2.0 fulfils the aspirations of the original hypertext pioneers, and the community that took up their work. However, it seems that the relationship between Web 2.0 and those original visions is more complex than this: many of the aspirations of the hypertext community have been fulfilled in Web 2.0, but as a collection of diverse applications, interoperating on top of a common Web platform (rather than as one engineered hypertext system).

Some aspirations are unsupported because they seem to be unnecessary for a given domain – so for example, Wiki pages are versioned, but Blogs are not. This indicates that versioning is chosen carefully, to avoid a user overhead if it is not really needed. Other aspirations, such as typed, n-ary links, are hardly supported at all. In the systems that we have looked at, this may be because of the way in which users are supposed to make loose, opportunistic connections, rather than to build a structured network. However it may be because hypertext is not the chosen format for such semantic networks, with system designers choosing to model their information behind the scenes in databases or ontological knowledge bases, and then translating to hypertext where needed (although Semantic Wikis such as Platypus[5] are challenging this trend).

Web 2.0 (meaning the set of applications, web sites and companies that define it) is not totally analogous to the vision of the early hypertext pioneers, mainly because the attributes that they were seeking are not available ubiquitously across all the systems of the Web. It is almost as if Web 2.0 has purposely rejected some of those old aspirations, and the assumptions that went with them, in favour of a more flexible, lightweight and responsive approach. It therefore offers an appealing updated or alternative vision, which addresses the same problems that those pioneers were attempting to solve. The Web 2.0 model is heterogeneous, ad-hoc, evolutionary rather than designed, but above all it is pragmatic and robust, allowing tools and applications to evolve naturally alongside each other, shaped by the communities that they serve.

6. REFERENCES

[1] Anderson, K. M., Taylor, R. N., and Whitehead, E. J. 1994. Chimera: hypertext for heterogeneous software environments. ACM Hypertext ‘94, pg 94-107.

[2] B. Christensen, G., Hansen, F. A., and Bouvin, N. O. 2003. Xspect: bridging open hypermedia and XLink. In ACM World Wide Web ‘03, Budapest, 2003) pg 490-499

[3] Bernstein, M. 1998.Patterns of hypertext. In ACM Hypertext ‘98, Pittsburgh, pg 21-29.

[4] Blackburn, S. and DeRoure, D. 1998. A tool for content based navigation of music. *ACM Multimedia‘98*,361-368.

[5] Brusilovsky, P., Eklund, J., and Schwarz, E. (1998) Web-based education for all: A tool for developing adaptive courseware. Computer Networks and ISDN Systems, 14-18 April 1998) 30 (1-7), 291-300.

[6] Vannevar Bush. As We May Think. *The Atlantic Monthly July ’45*. July 1945.

[7] George H. Collier, Thoth-II: hypertext with explicit semantics, ACM Hypertext ‘87, 269-289, Nov. 1987,

[8] Davis, H. C., Millard, D. E., Reich, S., Bouvin, N., Grønbæk, K., Nürnberg, P. J., Sloth, L., Wiil, U. K., and Anderson, K. 1999. Interoperability between hypermedia systems: the standardisation work of the OHSWG. In ACM Hypertext ‘99, 201-202

[9] De Bra P., Aroyo L., Chepegin V. The Next Big Thing: Adaptive Web-Based Systems, 2004, Journal of Digital Information, Vol. 5 Issue 1, Article No. 247, 2004-05-27

[10] El-Beltagy, S. R., Hall, W., De Roure, D., and Carr, L. 2001. Linking in context. In ACM Hypertext ’01. 151-160.

[11] Grønbæk, K., Bouvin, N. O., and Sloth, L. 1997. Designing Dexter-based hypermedia services for the World Wide Web. In ACM Hypertext ‘97, pgs 146-156, 1997.

[12] Halasz, F. G. 1987. Reflections on NoteCards: seven issues for the next generation of hypermedia systems. In ACM Hypertext ’97. 345-365

[13] Hall, W., Hill, G., and Davis, H. 1993. The microcosm link service. In ACM Hypertext ‘93. pg 256-259, 1993.

[14] David L. Hicks, John J. Leggett, Peter J. Nürnberg, John L. Schnase. A Hypermedia Version Control Framework. ACM Transactions on Information Systems, Vol. 16, No. 2, April 1998, 127 – 160

[15] Kahan, J. and Koivunen, M. 2001. Annotea: an open RDF infrastructure for shared Web annotations. ACM World Wide Web ‘01, Hong Kong, pg 623-632.

[16] Millard, D. E., Gibbins, N. M., Michaelides, D. T., and Weal, M. J. 2005. Mind the semantic gap. In ACM Hypertext ’05, pg 54-62.

[17] Nelson T. Literary Machines, published by the author, 1982 (reprinted by Mindfull Press, ISBN 0-89347-062-7)

[18] Delisle, N. & Schwartz, M. Neptune: a hypertext system for CAD applications. Proceeding of the ACM SIGMOD ’86, Washington, D.C., May 28-30, 1986, 132-142.

[19] Nürnberg, P. J., Leggett, J. J., Schneider, E. R., and Schnase, J. L. 1996. Hypermedia operating systems: a new paradigm for computing. In ACM Hypertext ’96. 194-202

[20] O’Reilly, T. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, self published on www.oreilly.com, 09/30/2005

[21] Reich, S., Carr, L., De Roure, D., and Hall, W. 1999. Where have you been from here? Trails in hypertext systems. ACM Comput. Surv. 31, 4es (Dec. 1999), 11.

[22] Walker, J. 2005. Feral hypertext: when hypertext literature escapes control. In ACM Hypertext ’05. 46-53.

[5] Platypus: http://platypuswiki.sourceforge.net/