Web 2.0
Next Big Thing or Next Big Internet Bubble?

David Best
Lecture Web Information Systems
Technische Universiteit Eindhoven

January 11, 2006

Abstract

In the past months, *Web 2.0*floats as a buzzword through the World Wide Web, signifying something new, exciting, and promising. This report aims at capturing the major ideas that stand behind this term, presenting them in a comprehensible way, trying to summarize and draw conclusions out of the current hype.

Contents

1 Introduction **1**

2 History **1**
2.1 Web 2.0 . . . 1
2.2 Web 1.0 . . . 1

3 Characteristics **2**
3.1 Rich User Experience . . . 2
3.2 User Participation . . . 3
3.3 Dynamic Content . . . 3
3.4 Metadata . . . 4
3.5 Web Standards – Valid Markup . . . 4
3.6 Scalability, 'The Long Tail' . . . 4

4 Technologies **5**
4.1 AJAX . . . 5
4.1.1 Overview . . . 5
4.1.2 Code Examples . . . 6
4.1.3 Advantages . . . 7
4.1.4 Disadvantages . . . 7
4.1.5 Example AJAX Applications . . . 8
4.2 Syndication . . . 8
4.3 RDF . . . 9

5 Applications **9**
5.1 Google . . . 10
5.1.1 Gmail . . . 10
5.1.2 Google Maps . . . 10
5.1.3 Personalized Google Search Page . . . 11
5.1.4 Google AdSense . . . 11
5.2 Microsoft . . . 12
5.2.1 Windows Live . . . 12
5.2.2 Office Live . . . 12
5.2.3 Snarf . . . 12
5.3 Blogs . . . 13
5.4 Wikipedia . . . 13
5.5 Flickr . . . 13
5.6 Flock . . . 14
5.7 Peer-to-Peer . . . 14

6 Semantic Web **15**
6.1 Overview . . . 15
6.2 Example . . . 15
6.3 Comparison to Web 2.0 . . . 15

7 Significance and Outlook **16**
7.1 User Participation 16
7.1.1 New Amateurism 16
7.2 Rich Web Applications 17
7.3 Relevance for Web Information Systems 17
7.4 Web 2.0 Hype 18

References **20**

1 Introduction

The idea that the World Wide Web is in a state of significant change, denoted by the term *Web 2.0*, is currently being extensively discussed. *Web 2.0*is hot, controversy, and maybe the (Internet world) buzzword of 2005.

We are going to examine what is being meant with *Web 2.0*, meaning why there is huge agreement on the fact that present changes have something in common which is different or new. We will identify this by examining the technologies and applications said to be *Web 2.0*.

One technology that was of major influence, and is of major significance for *Web 2.0*, is called AJAX and will thus be examined in a particularly detailed way.

Finally, we will try to capture the significance of *Web 2.0*: How successful are the new technologies and applications? How are major players on the Web reacting? To what extent are they already involved? Can there be a prediction made about how things will continue to change in the Web?

But to start, we will concentrate on a contrastive view on *Web 2.0*and what it is often compared to, the so-called *Web 1.0*.

2 History

2.1 Web 2.0

The term *Web 2.0*was created by the publisher Tim O'Reilly together with work partners in October 2004, as various sources concordantly state, amongst them O'Reilly himself [O'R05]. They then also held the first *Web 2.0*Conference, with its successor in this autumn.

In the past months, *Web 2.0*has been a hot topic for discussion, especially in the blogosphere[1], and this has also lead to first coverage in the so-called traditional media.

It is meant to summarize new technologies, applications, concepts, ideas and realizations in the World Wide Web. These vary from author to author; the term of *Web 2.0*is not a well-defined one, and we can only examine a selection here.

This term consists of the word *Web* and the version number *2.0*, so this implies that there exists at least one earlier version as a predecessor – the so-called *Web 1.0*.

2.2 Web 1.0

Often, *Web 2.0*is contrasted to what is considered *Web 1.0*, the traditional World Wide Web applications. Very popular is Tim O'Reilly's Web contrasting with the help of examples [O'R05]:

[1] the community of weblogging people, or more general, the set of weblogs

Web 1.0		**Web 2.0**
DoubleClick	→	Google AdSense
Ofoto	→	Flickr
Akamai	→	BitTorrent
mp3.com	→	Napster
Britannica Online	→	Wikipedia
personal websites	→	blogging
evite	→	upcoming.org and EVDB
domain name speculation	→	search engine optimization
page views	→	cost per click
screen scraping	→	web services
publishing	→	participation
content management systems	→	wikis
directories (taxonomy)	→	tagging ('folksonomy')
stickiness	→	syndication

The left-hand terms are often considered as unsuccessful and generally outdated, carrying a negative notion.

These contrastive examples refer to some of the characteristics and technologies that will follow in the next two sections.

In order to reflect the fuzziness of the term, we may have a look at another 'cloud of associations' before going into detail:

Image source: `http://www.andybudd.com/dcontruct05/images/meme3.jpg`, from [Bud05].

3 Characteristics

3.1 Rich User Experience

This term relates to the behaviour of a Web application. Many new *Web 2.0* applications are said to allow Rich User Experience because they behave in a way very similar to desktop applications. These latter are characterized through a state based user

interaction. This means that the application can have differing states, show progress, and reply fast, e.g. to bad input.

Web pages instead are based on the *HTTP* protocol, which does not support states.[2] For Web pages, this means that after each user action, a new page has to be generated, in order to display changes, validation, progress, etc., meaning the 'new state'. Many *Web 2.0*applications have overcome this by the use of the AJAX technology. More in section 4.1.

3.2 User Participation

The Web has changed when looking at the roles of providers and its users. In its beginning, it was a mostly single direction medium. Establishment of bidirectional use began when user participation in form of feedback possibilities like e-mail (`mailto:`) links or contact forms arose. Furthermore, message boards and guest books allowed the visitor of a Web site to comment and discuss.

The phenomenon of Social Networking [Wik05b] in the Web started with the growing popularity of Virtual Communities, generally bounded to one specific host and often based on shared special interests of the users[3]. Here, user profiling, messaging, chatting, discussing, etc. established rapidly in the Web.

With the idea of Wikis, user participation started into a new dimension: Users were not only allowed for the above in a restricted space, but also to create, modify, and delete the content of whole sites. Misuse is thought to be prevented by fast correction through the community with the help of kept previous versions of content.

3.3 Dynamic Content

Early Web sites were lay out to be static. Examples for this include personal home pages and corporate Web sites.

Dynamic content was introduced when the development of database back-upped Web sites began to increase. The databases were accessed by HTML-extending scripts written in languages such as CGI, Perl, PHP, etc. Example applications for this are online shops, virtual community platforms, and search engines.

When more and more personal home pages were becoming extended or replaced by weblogs, dynamic content became a design issue:

- Every entry is uniquely identifiable through a URI[4] (the notion of 'Permalink' – permanent links).
- Old content will not be deleted, but is still available through archives and/or permalinks (called 'trackbacks' in the blogosphere when linked to).

[2] In contrast to the *FTP* protocol, for example, where sessions with well-defined states are set up with every connection, possibly also anonymous sessions when an FTP server is supposed to act as an anonymous download server.

[3] we should not forget that Virtual Communities were formed in Newsgroups, the so-called Usenet, already before the Web existed

[4] Uniform Resource Identifier, a generalization of the URL which is restricted to the Web

This is also a main issue of the old alternative hypertext system *Xanadu*, which requires that 'every document is uniquely and securely identified' [Xan02]. This means that identifiers such as URLs will never get invalid.

Weblogs also implemented the following characteristic.

3.4 Metadata

In weblogs, it is common that articles are either 'tagged', meaning given one or more keywords (tags) about them, or put into categories. This belongs to the idea of semantic enrichment, creation of machine readable semantic data, the major issue of the Semantic Web. For a contrastive analysis of *Web 2.0*to the Semantic Web, please refer to chapter 6.

3.5 Web Standards – Valid Markup

Web standards are continuously being developed by the W3C, the World Wide Web Consortium. When finally released, they are expressed as so-called *Recommendations*.

Web pages that follow a syntax expressed in a W3C Recommendation are called *valid* and can be tested for validity at the W3C [W3C05a]. Validity of Web pages is a key feature often mentioned in combination with *Web 2.0*, not only in [O'R05].

But how looks reality in *Web 2.0*? One weblog author tested 34 often mentioned *Web 2.0*sites for validity, and only two passed without errors [Len05].

So the *Web 2.0*hype seems not yet to have triggered a new Web standards consciousness.

Still, the AJAX technology (see section 4.1) makes use of the open formats of the XML markup language and the Document Object Modelling (DOM) trees, being both W3C recommendations and thus differing from the proprietary, closed format of Flash, which provides a similar functionality.

3.6 Scalability, 'The Long Tail'

The term 'The Long Tail' refers to one property that is often attributed to Web success stories, such as peer-to-peer based applications, Amazon, Ebay, or Google's AdSense[5]. It covers the ability of applications to scale easily, and even to be of greater value the more users there are.

Peer-to-peer networks and Google's services are covered in detail in section 5, so let us have a closer look at Amazon's and Ebay's success.

- **Amazon** soon introduced user interaction, such as user-written reviews and their famous recommendation of other products of possible interest for a user, generated through buying statistics. These features scale easily with the number of their users, and are of course more significant, the more users there are.

[5]plus AdWords, see section 5.1.4

- **Ebay** relies on the principle of large supply and demand. The more offers there are, the more interesting for a customer the service becomes. The more prospective customers there are, the more attractive the service is for sellers.

These examples demonstrate a well-known general success formula which is of special significance for the Internet: the 'network effect'. This is an economic term, and it was there long before anyone talked about *Web 2.0*, or even the Web at all [Net05]. Probably one of the oldest examples for a network effect is a *telephone network*. Anyhow, Tim O'Reilly calls this in his famous essay 'a key Web 2.0 principle' [O'R05], and that is maybe why it was introduced as a *Web 2.0*feature.

4 Technologies

4.1 AJAX

4.1.1 Overview

AJAX is an acronym for Asynchronous JavaScript and XML. It is a technology that significantly changes the behaviour of Web pages:

- away from the traditionally necessary reload after a user action
- to a Web page that behaves in a way very similar to desktop applications.

Usually, a Web page has to completely reload after a user action, such as clicking on links, or entering data in a form. AJAX introduces an engine that runs on the client side – the Web browser. Certain actions can be carried out in the engine and need no data transfer to the server; they are carried out only on the client's computer and thus very fast, comparable to desktop applications.

AJAX is a combination of various techniques, of which one is JavaScript. Often, a huge amount of JavaScript code is loaded before any use of the Web page is possible, such as in Google's e-mail service Gmail. The Web page then already contains most of the data necessary for many actions, such as in Gmails case to browse through the e-mails, read, move, answer, delete them. Only some data may be necessary to reload, and this is being done through asynchronous Http requests with the use of XML based Web services like SOAP.

It is important to mention that AJAX is not a new technique – it just combines older ones. One which did not significantly strike attention among Web developers for years was *DHTML* (Dynamic HTML), providing similar functionality, only without the XMLHttpRequest of AJAX.

4.1.2 Code Examples

The JavaScript code is located in the HTML `<head>` of the page. Data retrieval from the server is being done via XMLHttpRequest, implemented by all major browsers in its current versions.

The following code examples are extracts of an online tutorial at [Ada05].

First, a JavaScript XMLHttpRequest object is being created (two methods differing for Internet Explorer and all other browsers):

```
var requester = new XMLHttpRequest();
var requester = new ActiveXObject("Microsoft.XMLHTTP");
```

Then, the asynchronous connection to the server can be opened and activated by

```
requester.open("GET","/feed.xml");
requester.send(null);
```

Data can then be sent as an argument of the `send()` method.

Through an event listener, the status of the XMLHttpRequest object can be checked, and if sending was successful, it can contain data in certain DOM[6] structured response properties.

User input data may be validated on the fly, for example by use of a JavaScript method to check a form. This method may be called when the user moves away from the form input field, which triggers a JavaScript `onchange` event.

```
function checkForm() {
  if (!receipt.valid) {
    receipt.focus();
    alert("Please enter a valid receipt number.");

    return false;
  }
}
```

This opens up a JavaScript pop-up alert window if the `valid` variable is set to `false`. The setting of this variable happens in a method that does the XML requests to the server.

[6] Document Object Model

4.1.3 Advantages

Web applications behave similar to desktop applications. Responses to user actions are very fast and the Web application resembles to have states although HTTP is a stateless protocol. Actions unusual for the Web, such as content changing without reloading, or drag-and-drop moving of elements, are possible now. This was only possible with the proprietary format Flash before, a product of Macromedia.

So, without the need for a browser plug-in, AJAX provides functionality similar to the ones that Flash based applications offer. But while these require a browser plug-in (installed by about 65 percent of Web users, see [Web05], AJAX does not. It builds on JavaScript, which 'only' about ten percent of Web users have disabled [The05], a significantly smaller number than with Flash.

4.1.4 Disadvantages

- **JavaScript** As mentioned, the AJAX engine is made up of JavaScript. Thus, currently ten percent of the Web users are excluded from AJAX applications.

  JavaScript has the expressive power of a complete programming language, in contrast to HTML, which is only a markup language. This means that Web applications in AJAX are much more error-prone concerning programming errors. They need extensive testing on various browsers and platforms to ensure that the desired functionality is present.

  Furthermore, JavaScript forms a barrier for Web access, an issue the W3C (World Wide Web Consortium) is concerned about Here, accessibility is defined such that 'people with disabilities can use the Web', but also 'people using a slow Internet connection', or 'people with changing abilities due to aging' [W3C05b].

  JavaScript is recommended to always be alternated with a `<noscript>` tag[7]. But if this is done, either the whole functionality of AJAX based Web pages must be re-written in a traditional manner, or is not available for about ten percent of the users that have disabled JavaScript (see previous section 4.1.3).

  In the Microsoft Internet Explorer, for users to have XMLHttpRequest to work, a major part of AJAX functionality, the active content module ActiveX must also be enabled 4.1.2. This excludes many users from AJAX use, as ActiveX is a known security problem in the Internet Explorer and thus disabled by many users.

- **The Back Button**

  Web browsers usually keep a history of whole Web pages. When a user presses the back button, he/she expects that the last action will be undone. This expectation does not conform with the effect that will appear with AJAX applications, as here single actions are not cacheable for the browser. Currently only workarounds to this problem are known, which complicate programming in AJAX.

[7] see `http://www.w3.org/WAI/wcag-curric/sam12-0.htm`

- **Bookmarking**

  As with many dynamically generated pages, bookmarking or linking to a certain state of such a page is nearly impossible, as these states are not uniquely identifiable by an URL. Solutions are combined with extra programming efforts, such as making use of URL fragment identifiers, i.e. anchors.

4.1.5 Example AJAX Applications

- **Google's New Services**

  In 2004 and 2005, the dominant search engine company Google introduced a number of new services, such as the e-mail service Gmail, the map service Google Maps, the personalized Google Search page, etc. All of these are based on AJAX technology.

  More on these in section 5.1.

- **The *Windows Live* Desktop**

  Microsoft recently introduced plans for and a beta version of a Web based, desktop-like application called 'Windows Live'. It strongly resembles Google's new personalized search page, including manageable pre-selected content, freely selectable content, or Web mail.

  More on this and the equally new *Office Live* service in section 5.2.

- **The Flickr Photo Service**

  This Web site uses AJAX for changes on a page such as a progress notification or picture title editing. More on Flickr in section 5.5.

4.2 Syndication

Overwhelmingly fast spread a technology called RSS, an acronym for either RDF Site Summary, or Really Simple Syndication. RSS is a well-defined syntax[8] for files in the Web that reflect updates on a Web site.

In the meantime, other formats for feeds have been developed and are in use, too, such as Atom[9], or Klip[10].

Examples for sites often offering syndication feeds are:

- Weblogs
- News sites
- Special topics (e.g. Web technology, cultural site, ...)
- Weather forecasts

[8] `http://blogs.law.harvard.edu/tech/rss`
[9] `http://www.atompub.org/`
[10] `http://www.serence.com/site.php?action=ser_products,devtools`

- Picture sites

This changes the way the Web user gets updates – from a pull-method to a push-method: The user does not have to go through the page to see if there are news, but the news are pushed to him as his RSS reader requests the feeds (the RSS files) regularly.

Of course, this kind of content syndication cannot only be used through stand-alone RSS readers, but also integrated into Web pages, such as the Weblog sidebar in MyOpera Community Blogs[11], or the Google personalized search page (see section 5.1.3). In the latter, syndication is also used as a mean of integration of multiple services.

It should not be omitted that personal news syndication was also present in earlier days of the Web. One example is the KNewsTicker for the Linux KDE window manager.

4.3 RDF

RDF (Resource Description Framework) is a specification to describe any type of resources. It is formally specified a s a sub-language of XML, but its data models can be expressed in various ways. Together with RDFS, the corresponding schema language, so-called ontologies for certain areas of resources can be modelled and formally described.

Two well-known RDF ontologies are:

- The vocabulary of the *Dublin Core Metadata Initiative*[12], which is, for example, often being used in HTML headers to extend the expressiveness of the HTML `<meta>` elements.
- The RSS syntax, which was formally expressed in RDF in earlier versions (see section 4.2 for more on RSS).

RDF(S) is supposed to be *the* language of the *Semantic Web*, but often mentioned in context with *Web 2.0*as well. More on the Semantic Web and a comparison to *Web 2.0*in section 6.

5 Applications

All of the following applications keep some of the characteristics that we have examined in the previous section. We will describe them shortly and try to find out about their relevance concerning *Web 2.0*.

[11] `http://my.opera.com/community/`

[12] see `http://dublincore.org/documents/dcmi-terms/` for an overview of the Dublin Core metadata terms

5.1 Google

5.1.1 Gmail

`http://gmail.google.com/`

Google's Web-based e-mail service Gmail was introduced in 2004.

Gmail is implemented with the AJAX technology, which implies that all the AJAX features mentioned in section 4.1 are present. Especially, Gmails interface resembles desktop applications, generating the so-called 'Rich User Experience' (3.1). Furthermore, many users with old browsers were excluded from Gmail a long time, until an alternative basic HTML version was made available in February 2005.

Among other features, Gmail allows without reload of whole pages to

- carry out all general mail program actions
- use keyboard shortcuts
- use an instant spell checker

which no Web mail site did before.

Gmail came with one interesting other feature, concerning the organization of messages. Unlike with probably all other major Web mail providers, it is not possible to organize messages into folders. Instead, users are meant to find mails by either

- **conversations** replies to and forwards of messages
- **starring** adding a 'star' to a message
- **searching** full text search on all messages, thus relying on Google's roots as being a search engine originally

One subject provoking huge discussion was the content related advertising in Gmail. This means that ads are displayed fitting the content of the text on the page respective page. For this, of course, mails have to be read by software, in order to determine which ad fits best. This feature also holds for Google's AdWords and AdSense products, but as e-mails are a more sensitive subject, it provoked debates and objections by many people.

5.1.2 Google Maps

`http://maps.google.com/`

Google Maps is a maps service providing satellite images of the whole world, together with street maps. It is included into 'Google Local', but was a stand-alone application a long time.

Two features are of special interest in the context of *Web 2.0*:

- **AJAX** This service is build on AJAX technology, too. This means, draggable satellite pictures, or navigation possibilities, that are very close to desktop applications, and quite fast responses also on slow computers and Internet connections.
- **Extensibility** Many Web services make use of Google Maps, for example the housing agency HousingMaps[13] for finding of accommodation via a Google Map, interwoven with data about houses from another Web site.

5.1.3 Personalized Google Search Page

This autumn, Google made its home page, `http://www.google.com/` and the respective localized pages, such as `http://www.google.nl/` and `http://www.google.co.uk/`, adaptable to a user's wishes. This concept is already known from providers such as Yahoo and msn. There, users could, after a login, choose from several modules, including news, weather, or e-mail. Google offers the same now, but offers the possibility to include almost any content, like RSS feeds or Web pages. Furthermore, the AJAX engine allows intuitive user actions like drag-and-drop, which is very unusual for Web sites.

5.1.4 Google AdSense

`https://www.google.com/adsense/`

Google AdSense is a service for webmasters to place advertisements on his Web site. These advertisements are context relevant – they are being retrieved from Google when a Web site is being visited by a user. The content that is displayed then depends on the content of the Web site[14]. The webmaster is paid per click, thus by popularity of his own site.

The corresponding program to place ads on other people's or businesses' Web sites is called AdWords. The price for an ad determines how often Google will display it, and whether it is being displayed on popular or less popular sites.

These ads were the first to really adapt to the page where they stand: they often are made up of not much more than some text, using the same styles as the Web site, and fit to it concerning the content. AdSense advertisements belong to the category of sponsored links (these are very frequent in search engines; often the first few search results are sponsored ones). These have proved to be more successful and better accepted than traditional, 'static' ads [Fit03].

So AdSense is a good example for the deep, extensive use of the power of the Web. Google combines its core area of excellence – indexing Web pages – with smart ideas and small scripts to generate a whole new field of income.

[13] `http://www.housingmaps.com/`

[14] Web sites that are below a certain level of popularity are provided with public service ads

5.2 Microsoft

5.2.1 Windows Live

`http://www.live.com/`

Windows Live is a platform to bundle certain user activities. After a login, it is possible for a user to create and modify a personal start page, very similar to Google's personalized search page (see section 5.1.3). A web e-mail interface, the *msn messenger* chat tool and a computer security centre are integrated into Windows Live, too.

The Web e-mail tool's behaviour is similar to the already longer existing Microsoft Exchange Web Mail platform. Here, the user had an intuitive and fast responding Web application, based on DHTML[15].

The conception of Windows Live signifies a reaction of Microsoft to Google's activities on the Web, especially with respect to binding users to a Web site by combining services and creating a user friendly and customizable environment.

5.2.2 Office Live

`http://www.microsoft.com/office/officelive/default.mspx`

Microsoft OfficeLive is a Web platform currently still before beta-status. It is targetted towards small businesses and allows them to set up a Web presence easily, with integration of e-mail addresses, and synchronizable with the desktop Microsoft Office programs.

Microsoft always aimed at integration of services and applications[16]. This platform is just another example for the growing attempt of Microsoft to do this on the Web, too.

5.2.3 Snarf

`http://research.microsoft.com/community/snarf/`

Snarf is an acronym for 'Social Network and Relationship Finder'. It was published on November 30, 2005, and is available for free download. It is an extension to Microsoft Outlook to organize unread e-mails by importance. Snarf displays them by importance. Criteria for importance are

- whether the mail is addressed personally, or through a distribution list
- metadata that Snarf collects about e-mail contacts, such as the number of e-mails received in a certain amount of time
- priorization of this metadata by the user.

[15] DHTML equals AJAX without XMLHttpRequest, see section 4.1

[16] one famous example is the integration of the Internet Explorer browser into the Windows operating system

Furthermore, Snarf allows displaying of e-mails in threads, meaning organized by replies, forwards, etc. This is a standard feature of Google's Gmail (section 5.1.1), but was offered by other e-mail clients long before, such as Mutt or KMail[17].

So this surely is not a Web application, but it introduces one interesting feature into e-mail management, which is the attempt to acquire meta-information about someone's contacts through analysis of their e-mail traffic, and a new presentation of e-mails resulting from this. Snarf makes use of the Social Network of someone's e-mail contacts, and it will surely prove whether it is successful, last but not least through a possible integration into future versions of the Microsoft Outlook e-mail client.

5.3 Blogs

Weblogs, or blogs for short, are online diaries. As they are visible to every visitor, its topics are mostly less personal than traditional diaries. Weblogs are characterized by hyperlinks to other articles or Web sites, a comment functionality, and unique identifiers for every entry and every comment. Many weblog platforms try to keep their applications Web standards conform, such as TypePad[18] and WordPress[19].

Blogging has introduced a whole new form of online publication, with huge success in the U.S.: In November 2004, 32 million people stated to read blogs regularly (about 27 percent of the Web users), while 8 million are authors (about 7 percent) [Hei05].

While blogging is a massive sign for growing user participation and interaction, some people criticize that it is often regarded as a replacement of the traditional media – see section 7.1.1 for details.

5.4 Wikipedia

`http://www.wikipedia.org/`

Wikipedia is an online encyclopedia, created and managed by its users. Editing pages is very easy, so the threshold for online collaboration to form a huge repository of knowledge is very low. Wikipedia is available in numerous languages worldwide.

Definitely, Wikipedia still suffers from lack of quality. This is an even bigger problem than the easily possible modification and deletion of content. Quality campaigns have been carried out, but still Wikipedia is not comparable in accuracy and reliability to existing encyclopedias.

Its main advantage is of course that it is free. Its authors form a community here, too, and the project must be considered a huge success in online collaboration.

5.5 Flickr

`http://www.flickr.com/`

[17] the Linux KDE Desktop Environment's e-mail client
[18] `http://www.sixapart.com/typepad/`
[19] `http://wordpress.org/`

Flickr is a photo publishing Web site. It was bought by Yahoo recently.

Flickr is the first photo service that introduced sorting photos by tags – i.e. keywords describing what is being depicted. Tagging is a concept of providing metadata described in section 3.4.

The Web site of Flickr makes use of AJAX technology, too. Moving images into and within sets can be done by drag-and-drop. Processes like uploading or renaming of images are supported by JavaScript actions showing progress and displaying changes immediately. In the beginning times, these used to be small Flash[20] elements, but that was changed soon.

5.6 Flock

`http://www.flock.com/`

Flock is a new Web browser that was recently released in a preview version. It is based on the freely available Firefox source code, thus resembling Firefox in many ways. What makes it special, is the integration of many services *Web 2.0*addicts will find handy:

- Facilities for direct blogging with four major Weblog services
- Bookmarking through the Web-based del.icio.us service
- Automatic RSS Feeds subscription for bookmarked sites
- A Flickr photo service plug-in

The Flock browser in fact equals Firefox together with many extensions, but the reason to make a whole new browser out of it is the simplification of the use of these services; the developers say that 'it should be simple for everyone to contribute to and participate on the Web'. Here we find one of the key *Web 2.0*properties again: user participation (see section 3.2).

5.7 Peer-to-Peer

Peer-to-peer applications function in a way opposed to the traditional client-server architecture. With peer-to-peer, every client is (meant to be) a server at the same time. This has the effect that with growing numbers of participants, not only the number of clients is growing (as in client-server architectures), but also the number of servers at the same time. Thus, they scale easily with growing success, as there is no need for additional server capacities.

Very popular are Internet file sharing applications, but the method is also used in academic environments, or in the BitTorrent network, a 'cooperative file distribution' implementation[21].

[20] Flash provides functionalities similar to AJAX; see section 4.1.3 for details

[21] see `http://www.bittorrent.com/introduction.html` for more information on BitTorrent

6 Semantic Web

6.1 Overview

The ides of a Semantic Web dates back to the end-1990s. It denotes an extension of the existing Web by machine readable metadata. In order to have a common syntax for this metadata, the description specification RDF has been developed. Here, facts are stored as triples of a subject, a predicate, and an object, such as 'John Doe is the webmaster of http://www.example.org'. Subject, predicate and object may be made uniquely identifiable through a URI[22], so that it can be used for other facts, when their denotion is clear for other fact authors.

6.2 Example

One example for the usefulness of metadata would be if images were supplied with a description of what they depict.

Let's imagine three images:

- Image No.1 depicting windows in a wall
- Image No.2 is a screen shot of a browser window
- Image No.3 depicts the Microsoft Windows logo

Searching for 'windows' in a search engine would probably return all three, although we might only have been interested in 'real windows', such as in image No.1. In the Semantic Web, these images would be described through a metadata file, say in the XML RDF serialization. Then each of them might have an RDF property like `depicts`, but a different object. These are then not distinguishable by name, but by their uniquely identifying URI.

If semantic data was present in large amounts, it would form a large graph – the Semantic Web. Here it would then, for example, be possible to carry out semantic queries, or to infer facts through following some graph's edges.

6.3 Comparison to Web 2.0

So how is the Semantic Web connected to *Web 2.0*? One *Web 2.0*characteristic is metadata as well – although not necessarily stored in RDF triples. In fact, RDF is not often heard of, or used, when it comes to *Web 2.0*and metadata. Furthermore, tags are a common means to express metadata. They are used in weblogs, but also in the bookmarking system del.icio.us, or in Flickr. A new term for classification through user collaboration is 'folksonomy', as opposed to the classical 'taxonomy', which means a structuring of things into a hierarchy. But this type of classification lacks the universality of the Semantic Web – which machine can tell if two tags of two different

[22]Uniform Resource Identifier, a generalization of the URL which is restricted to the Web

applications are the same or not? Even worse, the tag 'windows' still covers all of the three images above, without a possibility to distinguish them. So by tagging definitely a human made directory can nicely be created, and it scales perfectly with the number of tagged objects, which is a fact that makes them superiour to taxonomies. Still, expressiveness and quality of folksonomies are questionable.

After all it must be stated that the Semantic Web is a notion much more agreed on than *Web 2.0*– it is an older concept, and its properties and boundaries are quite clear. *Web 2.0*is still a kind of melting pot, a collection of items that often don't have very much in common.

7 Significance and Outlook

In this section, we will have a look at what key *Web 2.0*features really form something new in the World Wide Web, how they can be judged with respect to the development of the Web, and if a prediction for the future is possible.

7.1 User Participation

User Participation was present in the Web since its very beginning, and may be overestimated when speaking about *Web 2.0*. What is significant still, is that many projects and companies that are (still) successful today have always tried to make use of their users – Amazon, Ebay, Google, Wikipedia, but also peer-to-peer platforms[23]. These show that what Tim O'Reilly calls a *Web 2.0*principle still holds, namely that 'the service automatically gets better the more people use it'.

7.1.1 New Amateurism

This user participation and online collaboration is often praised as a new method of publishing, acquiring information, or categorizing. It must then compete with traditional ones, such as:

- Blogs vs. Print Media (and their online presence)
- Wikipedia vs. Encyclopedias
- Tagging vs. Editorial Classification

The main competitive advantages of the new are that they are free, and easily accessible. But often, they lack the quality of their traditional counterpart: Weblogs are very subjective, incomplete, and often simply wrong. Wikipedia is extensive, but not trustworthy. Articles from 'real encyclopedias' are written by experts, by highly skilled editors that are paid for their work, and whose work is based on decades of research. Many classifications, such as in scientific libraries, are structured in complex categories,

[23]from Napster to KaZaa, many shut down only for legal reasons

often very sophisticated, and developed over years. Tagging is highly intuitive, error-prone, and based only on keywords, not on any hierarchical structure between these.

Nicholas Carr discusses these and more issues of *Web 2.0* that are important to consider in his essay 'The amorality of Web 2.0' [Car05].

The main advantage, anyhow, of the new online collaboration results, is what Tim O'Reilly calls 'Harnessing Collective Intelligence' [O'R05]: making use of the power that arises when many people work together. Not only scale these systems easily with growing user numbers, but they get better as errors will be corrected through 'the mass': many eyes on every article in Wikipedia, so many fitting tags in Flickr that false tags will not matter when statistics and conclusions about numbers and connections of tags are drawn.

7.2 Rich Web Applications

One major part of *Web 2.0*, as should be clear from most of the example applications we examined, is the changed behaviour of Web applications: very many developers start to build sites that get closer to the user experience of a traditional desktop application. We might suppose that this may irritate current users of the Web, but surprisingly little contradiction has arisen – maybe also outweighed by the massive hype that arose around *Web 2.0*.

Important to notice here is the fact that this behaviour was already achievable through the use of Flash, but in reality, no major business Web site was completely programmed in Flash, by reasons of usability (see 4.1.3).

On the one hand, we must state that still the vast majority of Web sites are set up in the traditional way. But however, in more and more Web sites we find the use of AJAX, DHTML, or whatever technique used for creating desktop-similar Web sites.

The fact that the two companies that are definitely among the most relevant IT companies today, *Microsoft* and *Google*, are very busy building these new types of Web applications, is a sign that the hype is probably more than a bubble.

7.3 Relevance for Web Information Systems

As described in section 4.1, this new approach of Web application design makes use of HTTP requests. These exchange data between the application, which is running in the client, and the Web server. In case of AJAX, these are XML formatted data, but this is not necessarily the case. A sophisticated data management has always been a key issue of Web Information Systems research, so this kind of data management signifies a step into a direction away from the static, un-integrated data.

Metadata as in the idea of the Semantic Web still is not present in most of the *Web 2.0* issues, especially lacking universility and re-usability that could be achieved through the use of RDF. See section 6 for details.

Still, many *Web 2.0* success stories rely on making their applications open for further development, like extensions of Flickr, applications with GoogleMaps, or API's[24] of

[24] Application Programming Interface, an interface for programmers to a software system

Amazon, Ebay, Google, Yahoo. Often, these allow the use of Web Services; for an example see Amazon's list of Web Services at `http://www.amazon.com/webservices`.

7.4 Web 2.0 Hype

Definitely, the current discussion and expectations are very intensive and high, but temperature can be expected to cool down in the coming months.

New technologies are generally said to go through many phases – the so-called 'Hype Cycle', a term introduced in 1995 by a report of the GartnerGroup[25], 'When To Leap On The Hype Cycle', and appearing often in print media afterwards [Hyp05]. According to this Hype Cycle, new technologies go through five phases:

1. Technology trigger
2. Peak of inflated expectations
3. Trough of disillusionment
4. Slope of enlightenment / increasing realism
5. Plateau of productivity

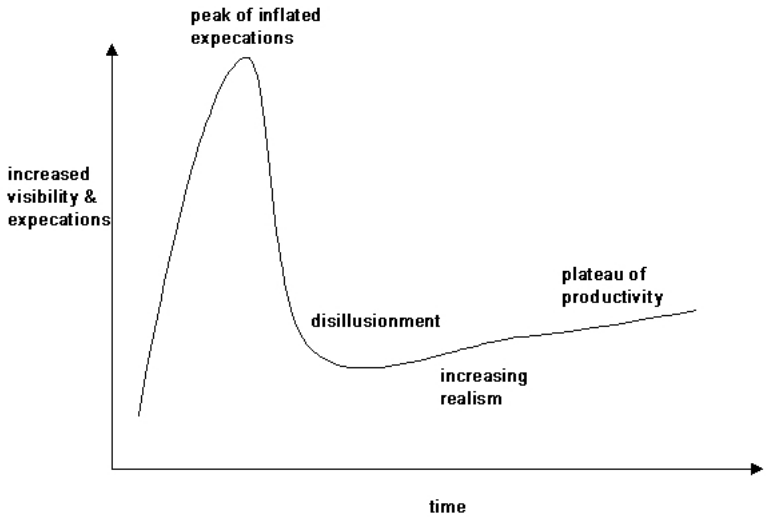

Image source: `http://www.riarlington.com/images/hypecyc.gif`

*Web 2.0*must here be located in phase two, the phase of 'overenthusiasm' where 'the only enterprises making money at this stage are conference organizers and magazine publishers' [Hyp05].

The question is, will *Web 2.0*really go through the following phases? Because of the diversity of its characteristics and applications, it may also be possible that people will

[25]'Research and analysis of information technology companies, products, and services, and of several industry sectors' (`http://www.gartner.com/`)

no longer be willing to use one single term to cover all these, and every technology will go its own way.

Predictions are always hard to make, especially in a so fast changing domain as the World Wide Web. This is the reason why we will not try to prophesy anything. What we can state, is the fact that there are significant changes, that there is a great hype, and that there are plenty of new ideas.

References

[Ada05] Adams, Cameron: *AJAX: Usable Interactivity with Remote Scripting*. Version: July 2005. http://www.sitepoint.com/article/remote-scripting-ajax. – Online–Resource, last checked: 2005-12-03

[Bud05] Budd, Andy: *What is Web 2.0?* Version: 2005. http://www.andybudd.com/dcontruct05/. – Online–Resource, last checked: 2005-12-03

[Car05] Carr, Nicholas: *The amorality of Web 2.0*. Version: October 2005. http://www.roughtype.com/archives/2005/10/the_amorality_o.php. – Online–Resource, last checked: 2005-12-04

[Fit03] Fittkau: *17. WWW-Benutzer-Analyse W3B*. Version: November 2003. http://www.wuv.de/daten/studien/112003/809/2610.html. – Online–Resource, last checked: 2005-12-02

[Hei05] *32 Millionen US-Amerikaner lesen Blogs*. Version: January 2005. http://www.heise.de/newsticker/meldung/54730/. – Online–Resource, last checked: 2005-12-03. – Heise Online News

[Hyp05] *Hype Cycle*. Version: December 2005. http://www.wordspy.com/words/hypecycle.asp. – Online–Resource, last checked: 2005-12-03. – A Word Spy Definition

[Len05] Lennartz, Sven: *Web 2.0 und Standards*. Version: October 2005. http://www.drweb.de/weblog/weblog/?p=428. – Online–Resource, last checked: 2005-12-03

[Net05] *Network Effect (Definition)*. Version: 2005. http://www.economist.com/research/Economics/alphabetic.cfm?LETTER=N. – Online–Resource, last checked: 2005-12-03. – Adapted from 'Essential Economist', published by Profile Books

[O'R05] O'Reilly, Tim: *What is Web 2.0*. Version: September 2005. http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html. – Online–Resource, last checked: 2005-12-03

[The05] *JavaScript Statistics*. Version: October 2005. http://www.thecounter.com/stats/2005/October/javas.php. – Online–Resource, last checked: 2005-12-03. – TheCounter.com Web Site Analysis, Global Statistics

[W3C05a] W3C: *Markup Validation Service*. Version: October 2005. http://validator.w3.org/. – Online–Resource, last checked: 2005-12-03. – W3C – World Wide Web Consortium

[W3C05b] W3C: *Web Accessibility Initiative (WAI)*. Version: November 2005. http://www.w3.org/WAI/. – Online–Resource, last checked: 2005-12-03. – W3C – World Wide Web Consortium

[Web05] *Web-Barometer*. Version: December 2005. `http://www.webhits.de/deutsch/index.shtml?webstats.html`. – Online–Resource, last checked: 2005-12-03

[Wik05a] WIKIPEDIA: *Ajax (Programmierung)*. Version: November 2005. `http://de.wikipedia.org/wiki/Ajax_(Programmierung)`. – Online–Resource, last checked: 2005-11-09

[Wik05b] WIKIPEDIA: *Social network*. Version: November 2005. `http://en.wikipedia.org/wiki/Social_network`. – Online–Resource, last checked: 2005-11-28

[Wik05c] WIKIPEDIA: *Web 2.0*. Version: October 2005. `http://de.wikipedia.org/Web2.0`. – Online–Resource, last checked: 2005-10-19

[Win05] *Windows Live Ideas*. Version: 2005. `http://ideas.live.com/whatis.aspx`. – Online–Resource, last checked: 2005-12-03

[Xan02] *Xanadu FAQ*. Version: April 2002. `http://www.xanadu.com.au/general/faq.html`. – Online–Resource, last checked: 2005-12-03