NetworkedPlanet

Web 3.0

Subjects On The Web

Prepared By:	Kal Ahmed
Prepared For:	
Version/Status:	
Date:	16 February 2009

1 Introduction

In the past 20 years, the Web has developed from a niche technology to a mass-media providing new forms of communication and interaction between people. Web 1.0 was a technical platform - a common set of protocols and formats that allowed machines to communicate and present information from a remote server to a local user. Web 2.0 used the technical platform of Web 1.0 to build more interactive web sites where users contribute and share content and become creators and owners of content rather than passive consumers.

Web 2.0 has reached the limits of what can be achieved on the technical platform of Web 1.0; new technologies must be put in place to provide a fundamentally new technical infrastructure, or platform, to enable the next generation of innovative web applications. Key to this Web 3.0 platform is a set of protocols and formats that allow the communication of subjects and people's perceptions of those subjects between computers, and that enable new applications to be built that allow users to create, share and integrate information and knowledge seamlessly.

The new Web platform will no longer be about using a browser window to retrieve information from one server and then from another server. Nor will this platform be based on portals or search engines that provide us with links to pages. The new Web will be based on applications that bring us relevant information from all across the Internet and bind it together for us, presenting us with our own personal view created from semantic structures taken from sources we trust and new sources we want to explore. This platform will also allow us to see authoritative content from trusted sources alongside commentary from our peers and will enable us to contribute to debate and form new social networks focussed around the subjects and semantic structures we are interested in.

Imagine an application that knows all about music. It knows about all the great composers and their works, it knows about the performances of those works and the recordings of them, it knows about every gig that the Beatles ever played and it knows about all the future gigs of all the Beatles tribute bands out there, including the ones in you area in the next month. Not only that but it is able to connect you to other classical-music-loving Beatles fans out there so that you can discuss your shared passions, and it is able to act as a channel, constantly receiving updates from all around the Web so that any time you return to that aspect of your life you can instantly see what is happening around the world in that area. It also acts as a broadcast channel on which our own comments, thoughts and new insights can be made available to anyone who is interested and which doesn't rely on artificial measures of popularity such as links in and out of pages.

In the rest of this paper we discuss the basic concepts of the Web 3.0 platform. Remember that we are talking about the technical platform for a new generation of web applications, we can only start to guess at the applications that could result from a combination of the platform we are proposing and the creative genius of web developers everywhere.

2 Subjects and Perceptions

In our vision of the Web 3.0 platform, the traditional web of pages of content and services is enhanced with a new layer of subjects and perceptions that allow people to more accurately communicate knowledge and information.

Put simply, a **subject** is anything at all that we want to talk about - cultural subjects such as a piece of music or a film; social subjects such as a person, an event or a relationship; business subjects such as a company, a department or a stock; even non-tangible subjects such as love and friendship. For a subject to be shared it requires only a definition (which may be an authoritative definition or a personal one) and an identifier. The identifier provides the "hook" for our computers to know when two people are discussing the same subject. Anyone can create an identifier for a subject and an **identity resolution service** is used by machines to determine when two different identifiers refer to the same thing.

Perceptions can be most simply thought of as snippets of fact or opinion that are encoded in such a way that a machine can determine what subject(s) the fact is about, and that enables it to relate that fact to other facts with minimal human interaction or guidance. A perception might encode the fact that Beethoven composed his 5th Symphony between 1804 and 1808. Another perception would record the premiere of Beethoven's 5th symphony took place on 22nd December 1808. A third that the premier of Beethoven's 5th symphony was conducted by Beethoven himself. A fourth that Beethoven's 5th symphony is available in recorded form from a particular retailer or that a new recording of the symphony is to be released next week. Like their more conventional namesake, perceptions are personal to us - we each have our own perspective on a subject and that shapes our perceptions about it. Discovering, discussing and challenging the perspectives that other people hold on the same subject is the core of learning and communicating.

A perception is web content, meaning that it has an address where it can be retrieved from and it can be transmitted, cached and transformed through all of the standard mechanisms of the current web. Just as text on the web is typically encoded in HTML to allow machies to process it, so perceptions require their own format. A number of formats already exist for encoding perceptions including standard formats such as RDF from the W3C and ISO Topic Maps. What makes a **perception** on the Web 3.0 platform different to a piece of text like the previous paragraph is that in a perception, the identity of the subject we are talking about (Beethoven, Beethoven's 5th Symphony, a date such as 22nd December 1808, or next week) is made explicit and made in a way that allows a machine to find other information related to the same thing. Furthermore the relationships expressed in a perception are also subjects, and so also use identifiers to define the types of relationships described by the perception. These differences mean that a perception is machine processable - a computer can interpret the content of a perception to determine if and how it should be displayed to the user.

A second function of the identity resolution service is therefore to find all perceptions relating to a given subject. Again the resolution service can make use of information about identifiers that are used to represent the same subject and because perceptions are web content the identity resolution service simply stores the links to the perceptions themselves. In this way the identity resolution service acts like a cross between a DNS server that resolves names to machine addresses and a search service that finds all content related to a query.

3 Linking Perceptions To Services and Content

While perceptions are effective carriers of information, it is highly unlikely (no matter how desirable) that all of the existing knowledge of the world could be converted to subjects and perceptions. The Web will still consist of content that is not expressed as perceptions and of services that cannot be easily represented in this way. For that reason perceptions must be able to reference both content and services that exist elsewhere on the web. The mechanisms for doing this are already built in to the basic Web 1.0 platform, so this is not a difficult thing to achieve - a perception can be used to encode a relationship between a subject and a piece of web content or a web service, but using an identifier to state how the subject and the content are related. For example a CD represented as a subject might have one perception that links to its cover art, and multiple perceptions that link to the shopping services of different vendors or to a price comparison site. This means that the perception that links subjects to content is independent of both the original subject definition and the linked content - allowing any person or organisation to effectively annotate both Web 1.0/2.0 content and Web 3.0 subjects using the same simple mechanism.

Furthermore it is possible that over time the facts and assertions contained in web content could be migrated to subjects and perceptions. This might be done manually for example by volunteers reading through Wikipedia pages and creating subjects and perceptions for what they discover there; or it might be done automatically using textual analysis tools.

4 Publishing Concepts and Perceptions

While the Identity Resolution Service is one part of the Web 3.0 platform there is also a need for tools that provide individuals and groups a means not only to publish subjects and their identifiers but also to create perceptions of subjects.

In the same way that a Wiki application enables anyone to create human readable content on the web so there will be a class of application that easily allows users to create and publis structured semantic information.

These tools and services will be provided not by any single vendor or product but will be stand alone applications, web services and extensions to existing publishing platforms. The identity resolution service ensures that however these perceptions are structured, that they can be brought together by a myriad of different clients.

Applications that support the authoring of perceptions will need to allow an author to create relationships with other perceptions that are published on the web. These relationships will not be undefined as is the case with the HTML links of the Web 1.0/2.0 platform, but will be typed and named such that they take on more significance for end users. These same end users will then be able to navigate a web of structured data and not just hyperlinked text.

5 Scenario

We do not imagine that Web 3.0 desktop applications will be limited to browsers, all kinds of internet-connected applications will be able to make use of this platform. Here is just one scenario.

Adam is playing a new recording of Bach's Cello Suites through his web 3.0 enabled media center application. The application uses meta tags in the MP3 files to get an identifier for the composer (in this case J.S Bach), and the piece (Cello Suite #2 in D Minor) and sends those identifiers to an identity resolution service to find what perceptions are available. Adam is a keen music shopper and has already set his preferences to indicate which services he prefers for the download of music. His browsing history also shows a keen interest in biography both in the web sites he visits and the books that he reads, indeed he has a number of biographers bookmarked as being subjects he is interested in hearing about when it is relevant.

The search on JS Bach returrns many perceptions, but through Adam's preferences for biography and music and book related purchases the application can filter out perceptions that provide opinion of film or video resources. Biographical entries such as in an online Grove or on Wikipedia are promoted, as are perceptions relating to publications that are about JS Bach.

After the first set of queries are executed, perhaps in the background as Adam starts to browse one of the recommended biography web sites, the web 3.0 application expands the first set of results by querying for perceptions related to the most interesting subjects returned in the first set of results. This query discovers that one of the publications about JS Bach that has recently been released is by one of Adam's bookmarked favourite subjects. As this is a subject Adam has asked to be notified about, the application pops this information up immediately on discovery allowing Adam the choice to bookmark that publication subject for exploration later or to ask the web 3.0 application to find all perceptions that provide purchasing information for the book.

This combination of exploration - both automated and human guided - provides a deeper and richer experience that no longer has to be purely a "Web" experience but can be easily integrated into all kinds of applications to make the wealth of information available on the Internet a part of all kinds of daily tasks.

6 Summary

Web 2.0, built on the existing web platform of pages of content and links has reached the limit of what it is capable of. The prinipal limitation is the difficulty faced in findng new sources of relevant content in a web of pages and links.

The Web 3.0 platform adds the notions of subjects and perceptions about subjects as a new additional layer in the web stack. Subjects allow people to express explicitly what they are talking about and perceptions allow them to encode facts and opinion in a way that allows them to be compared to and combined with the facts and opinions provided by others. In this way the Web 3.0 platform enables new social networks to be built around subjects of common interest.

Web 3.0 is not a replacement for the current Web. It is another incremental step towards the goal of improving communication, social connection and knowledge across the world.