



The Oxford Times

Nov 24 2008

LATEST HEADLINES

Language meets maths

4:03pm Thursday 20th November 2008

The constantly evolving nature of the Internet means many businesses are left clinging on to the coat tails of major advances. Web 2.0 — not a new version of the Internet but new ways of using it, including social networking sites and the use of videos and blogs — has now been embraced by many.

But now IT professionals are starting to talk about Web 3.0 which is exploring a new range of possibilities yet to be developed.

Web 3.0 actually describes a paradigm shift in how information is structured and searched for on the World Wide Web.

Web 3.0 concerns the use of the Semantic Web — it's about enabling search engines to scan for meaning and interpretation when presented with a search query, rather than just corresponding the number and density of matching search terms.

The Semantic Web is the brainchild of Tim Berners-Lee, creator of the World Wide Web and head of the World Wide Web consortium, the only independent organisation dedicated to consolidating standards of quality in content and structure to the web.

The repercussions of this shift towards a more intelligent and intuitive web will have a huge effect on all web-based data, be it commercial, academic or cultural.

Websites will either have to adapt to the Semantic Web, or be left outside the search loop.

According to Lyang Yu, author of The Semantic Web and Semantic Web Services, most of us use the web for three basic functions — search, integration and data mining, in that order.

The current search paradigm is epitomised by Google — answering particular queries by matching documents from its database that correspond to the precise language the searcher has entered.

Search Spiders are like stupid robots that only do precisely what you tell them. As a result, most search engine optimisation (SEO), tactics are largely restricted to optimising the use of commercially strategic language on websites that will more likely be coughed up by Google or Yahoo in response to a search query.

Google's AdWords operates on the principle of auctioning off the most popular searched for vocabulary to the highest bidder.

This year, the growth in search engine optimisation activities by brand leaders is unprecedented and rapidly increasing. Any blue chip worth its salt has an in-house search engine optimiser at the centre of its web development team.

Big brand companies like Tesco and Renault are spending their Internet budgets on increasingly competitive SEO tactics, both on page and off page.

One of the biggest growths in IT recruitment in both the US and the UK is for SEO managers and experts. Some companies currently spend as much optimising their web presence as they do designing their sites.

Usability and optimisation have become the defining criteria for distinguishing an effective website from a rarely visited online brochure.

While keyword search is still the most popular search method, it is seldom accurate. Users sometimes get up to 10,000 hits on a result page and then have to wade through a list of loosely-related keyword results to find the relevant documents they were searching for.

Up until the Semantic Web, search criteria have been based on the choice of the correct key words to tag and identify your web presence to the spiders indexing the Internet.

The premise is that the closer the language of your website corresponds to the language choice of the searcher, the more relevant your website, and the more likely Google will rank you higher in the search results.

But this is all coming to an end. The Semantic Web is where language finally meets mathematics.

In comparison to standard search, semantic search looks at the logic of the sentence — how words in a sentence relate to one another, as well as understanding the context of the keywords.

Instead of clumsy, corresponding criteria, the words grouped around a keyword or phrase will now play as important a factor in the relevancy of the term as the keyword itself. The focus is now on context, how words and assets are grouped together.

For example, when a term is ambiguous, such as with the word bark, semantic analysis is needed on the other words that wrap around it to give it its true meaning and context.

So a semantic web search for Obama plus McCain would correctly interpret that the searcher was seeking results relevant to the recent election campaign, as opposed to results that contained those names.

There are search engines on the net that are already beginning to harness the principles of semantic web development.

Cuil.com is a semantically inspired search engine that pulls relevant results from deep within website pages as opposed to just listing the index page of a particular website.

Other good examples include juiceapp.com, cuil.com, illuimin8.com and headup.com.

Ultimately web content publishers are going to have to adapt to the notion that all published content is equally accessible to semantic-based search engines. The past SEO criteria for priority placement of text and other assets on a website is declining.

Yahoo's recently launched SearchMonkey applications are semantically based search tools. Yelp, Yahoo!Local, and LinkedIn Enhanced now appear automatically in Yahoo search results.

These three applications are among the first to share structured data. ZIMesh is a semantic information management and recommendation engine that manages

personal information. ZIMesh is powered by a semantic platform, which automatically tries to understand users' interests over time, and connects them to topics, users and contents it thinks will be of interest.

So what does this shift represent to the average business trying to maintain a competitive web presence?

In one way it means that the playing field is levelling. There is less need to spend a fortune on SEO tactics that are never guaranteed to deliver measurable results.

Writing concise, relevant and informative content for a website will always pay dividends over SEO tactics. Google wants to refer its queries to useful resources. The more useful a resource you make your website, the more likely Google will rank you highly. So, do nothing but be good.

The more the individual data components — text, images, video, sound — of your website are ascribed searchable terms, the larger and deeper your web presence and the more likely you will be found by a semantic search.

The end result will be a more responsive, more intelligent and ultimately more useful world wide web.

o Contact: Igor Goldkind, 01865 201955 Web: www.oxfordseo.com