

Web of entities: prepare to 'Okkamise'!

March 7 2008

Internet searching is something of an art form. The spaghetti-like tangle of documents and fragments resulting from what you thought were perfectly cogent keyword searches make the web a forbidding place. European researchers are developing a better way to publish, link and find information using a “web of entities”. Prepare to Okkamise!

That the word “Google” has entered our vocabulary with such ease is testament to the powerful yet complex “web of documents” that we call the internet. If I want the number of a nearby trattoria in Brussels, but I can't remember its name, I enter the keywords “trattoria and brussels” and ecco the results are displayed... all 25,000 of them! How am I supposed to find the restaurant I'm looking for? More searching, more hassles.

In some ways, the internet's success threatens to undermine its ultimate utility unless a better way to structure the information is developed. This is where Okkam enters the picture.

Less hassle, better searches

The idea behind the EU-funded Okkam project is to unlock the full potential of the semantic web, helping people and machines to find, share and integrate information more easily. It borrows from ‘Ockham's razor’, a principle named after 14th-century logician William of Ockham that assumes the simplest solution is the best. “Entities should not be multiplied beyond necessity,” it states.

With Okkam, the main ‘objects’ being scanned are no longer documents that just happen to contain certain keywords, but ‘entities’, such as people, locations, organisations or events, explains Paolo Bouquet of Trento University and Okkam’s spiritual leader.

The core Okkam infrastructure will store and make available for reuse so-called “global identifiers” which can be applied to and used by anyone or anything across formats and applications. These are not to be confused with “certification”, stresses Bouquet, which he says are targeted more at making the web a safer place to transact. It is more concerned with distributed information and knowledge management.

Big companies, for example, can quickly and accurately benchmark their new products or processes against competitors or carry out internal knowledge management tasks. Project partner SAP, the enterprise software giant, is testing how Okkam can help in managing information on their public web portals like sdn.sap.com. Other Okkam partners, the scientific publisher Elsevier and ANSA, Italy’s leading news agency, are defining the authoring environment for scholarly and news content, respectively.

“One of the biggest risks we face,” Bouquet tells ICT Results, is people thinking the identifiers are a controlling device, a ‘Big Brother’ scenario. Far from it, the information that we (and you as an Okkam user) gather is the bare minimum to improve web searches. So you can quickly discern, for example, whether ‘Paris’ is the capital of France or a bistro in Boston, and whether it’s a web-page or an obscure mention in a Voltaire manuscript.”

Future networks, present challenges

Trends in the semantic web and social networking are ushering in a new era of meaningful and mobile information searching and interaction

online.

The future network is moving away from people sitting at home in front of their PCs trying to find information in billions of unstructured pages using what Bouquet calls “keyword guessing”.

Okkam’s coordinator says more precision and integration are inevitable developments on the net: “Information will be integrated and clustered from a large number of different, heterogeneous data sources all over the internet, provided by software agents, responding to users' data needs in whatever contexts.”

Of course, this scenario calls for a serious rethink of the ‘publish and be damned’ approach to Web 1.0 and even 2.0. “We believe that Okkam represents a substantial move in the direction of a ‘web of entities’,” he posits. But are we ready for this interpretation of Web 3.0?

Okkam’s entity identifiers offer a powerful departure on the traditional online social networking scene, where you post bits and pieces about yourself on, say, LinkedIn then some more on Flickr. What happens to the data then? Can it be corralled together? Bouquet thinks so.

For example, he says, with their Foaf-O-Matic application you can generate “Foaf profiles” using Okkam infrastructure to issue friends with globally unique identifiers which can be used on multiple social networking platforms creating one big “distributed and decentralised social network”. But having the technical ability and making it happen is not the same thing, he concedes.

Huge but perception is key

William Stevens of Europe Unlimited, a networking consultancy, echoes this view. The technology for the semantic web is sound, but getting

people to perceive the new developments as useful is the trick, he suggests. “Not just an attractive technology, but one that’s actually used on the market.” A critical mass of users and entries will mean the difference between useful Okkam searching and lacklustre results, he notes.

Although very early days for the project, the plan by the end of 2008 is to have a solid starter-base of 1 million ‘entities’, with a further million every year for the duration of the 30-month project. Without this critical mass, it will be harder to convince early adopters, especially application developers to take up and use Okkam.

“Having a solid business exploitation model and sustainability strategy built into the research plan is also critical,” Stevens says. It will also be important to get the word out about Okkam to application developers, industry, investors and users the world over.

Having SAP and Elsevier, two big potential users, as partners is clearly no coincidence. Bouquet has also presented demos of the technology and Okkam’s business plan to big names in the business, including Sisco and Microsoft, at the recent i-techpartner Software Forum in Porto, hosted by Europe Unlimited.

“Okkam definitely generated some buzz in Portugal!” confirms Stevens.

More information: www.okkam.org/

Source: [ICT Results](#)

Citation: Web of entities: prepare to 'Okkamise'! (2008, March 7) retrieved 26 April 2024 from <https://phys.org/news/2008-03-web-entities-okkamise.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.