Web & lib 2: a sceptic’s view

By Thomas Krichel

Born in Völklingen, Sarre, in 1965, studied economics and social sciences in the universities of Toulouse, Paris, Exeter and Leicester. Between 1993 and 2001 taught at the Department of Economics of the University of Surrey. In 1993 fundó NetEc, a consortium of projects on internet for economists university, and in 1997 the repository RePEc of documents on economics. In 2000 was visiting professor at the Hitotsubashi University, Japan. Since 2001 teaches in the College of Information and Computer Science of the Long Island University. Since July 2005 supervises projects of students in the Faculty of Information Technology of Novosibirsk State University, Siberia.

“C’EST L’HOMME QUE JE SUIS QUI ME REND MISANTHROPE”, noted Jules Renard. Yes, I don’t care much for the concepts of Web 2.0 and Library 2.0 and other buzz words a la mode. Call me a dinosaur, but I don’t blog and don’t read blogs. But even cynics have to admit that the Internet continues to transform information services like nothing ever before. Some innovations will stay, others will be gone in a few years.

The beginning of the millennium has been the moment of the search engine, especially for Google. Here we talk about the big machine that soaks up all the pages on the web. Its result is quite frightening. I have a farm of web servers. From casual observation of my web logs, over 50% of all accesses are now done by robots working on behalf of search engines, and the remaining part of usage seems to have come from users who have found my contents through search engines. The dependency is quite worrying. In olden days, people linked to your site, if it was any good. These times seem to have gone. Welcome to machine. Machine reads all the web, users only ask machine. What machine does not know about, nobody can see.

As to the second part of the first decade, some of the interest seems to have shifted. Machine Google has not been innovating much in the past couple of years, the search seems pretty much to be the same, and on a day-by-day experience, it seems to be missing an ever larger part of the answers. I can easy find documents that ought to be found but don’t seem to be there. In the meantime, interest has shifted to something quite different. Away from machine a number of grass-roots up-shoots have come along to provide services to smaller communities or that are run by small communities. Thus a new empowerment of users seems to be taking place. While none is powerful overall, such start-ups are interesting to the information professional that needs to look new ways to deliver information and at new ways to get users involved.

Let me take you for a quick stroll over what’s written here in this issue. I’ll start from the back. The blog is commonly thought of belonging to Web 2.0 and Library 2.0 and other buzz words a la mode. Call me a dinosaur, but I don’t blog and don’t read blogs. But even cynics have to admit that the Internet continues to transform information services like nothing ever before. Some innovations will stay, others will be gone in a few years.

The beginning of the millennium has been the moment of the search engine, especially for Google. Here we talk about the big machine that soaks up all the pages on the web. Its result is quite frightening. I have a farm of web servers. From casual observation of my web logs, over 50% of all accesses are now done by robots working on behalf of search engines, and the remaining part of usage seems to have come from users who have found my contents through search engines. The dependency is quite worrying. In olden days, people linked to your site, if it was any good. These times seem to have gone. Welcome to machine. Machine reads all the web, users only ask machine. What machine does not know about, nobody can see.

As to the second part of the first decade, some of the interest seems to have shifted. Machine Google has not been innovating much in the past couple of years, the search seems pretty much to be the same, and on a day-by-day experience, it seems to be missing an ever larger part of the answers. I can easy find documents that ought to be found but don’t seem to be there. In the meantime, interest has shifted to something quite different. Away from machine a number of grass-roots up-shoots have come along to provide services to smaller communities or that are run by small communities. Thus a new empowerment of users seems to be taking place. While none is powerful overall, such start-ups are interesting to the information professional that needs to look new ways to deliver information and at new ways to get users involved.

Let me take you for a quick stroll over what’s written here in this issue. I’ll start from the back. The blog is commonly thought of belonging to Web 2.0. To me, a blog is just a tool that allows a technologically challenged person to operate a web site. These folks don’t need to study html, CSS and usability they just do it. “A collaborative LIS weblog: details of the planning and development of Sedic’s blog” is a paper by “el equipo de redacción del blog de Sedic”. I first thought the Sedic was something to eat, but a Google search revealed that it is actually the admirable Asociación Española de Documentación e Información. On its pages, it proclaims the profound insight that “SEDIC-blog es el blog de SEDIC”. As an “official” blog, quality has to be ensured in some ways. Participants are segregated by type, and the blog is subdivided into topic areas. There has to be a balance struck between the wish to attract a wide range of collaborators engaging in free, spontaneous discussion, versus the requirement of professional style and classificability of contribution. The Sedic blog gang does illustrate this well.

http://blog.sedic.es/

In the paper that precedes it, called “365 days of books”: a blog to promote reading’ Francisco López Hernández and Honorio Penadés de la Cruz discuss a blog that serves the promotion of literacy to the university population. I am a bit skeptical that university
people need to read more in their leisure time, when really they can not cope with the professional literature. But upon visiting the blog I was impressed by the variety of books discussed, as well as by the visually attractive design. The article intelligently relates the creation of the blog with the work of the library, and the situation of the information profession in general. The authors explicitly refer to the work of the Sedic. It seems to have been an inspiration to others.

http://365diasdelibros.blogspot.com/

The paper “Evaluación formal de blogs con contenidos académicos y de investigación en el área de documentación” by Sonia Jiménez Hidalgo y Javier Salvador Bruna enters more formal academic terrain. The wording of the abstract seems unfortunate. It is simply not true that blogs “are tools for editing and publishing any kind of contents in the Internet”. You can not, for example, edit a sound file using a blog. Fortunately, the authors come later to their senses and implicitly say they “estamos hablando de un portal web más o menos interactivo que publica breves reseñas comentadas y que es actualizado de forma frecuente.”. They have a bunch of criteria for evaluation, but none of them pertain to the subject matter of the blog itself. The authors provide empirical data on their findings about blogs that are mildly interesting.

“Creo que promover el uso de Dublin Core para mejorar la recuperación de información en la Web es como espolear un caballo muerto”

Much to our relief, the previous paper, by comrades Eva Méndez, Alejandro Bravo and Leandro Mariano López, is not about blogs, but about “Microformats: web 2.0 for Dublin Core”. In my opinion his title is a little confusing and perhaps there should be some more explanations but it is very interesting. I personally find that promoting the use of Dublin Core to improve information retrieval on the web is flogging a dead horse. Microformats can help if they provide a syntactic basis for strings that are sufficiently rare otherwise that they can be used as informal identifiers. While most people will be familiar with Dublin Core, they may not be familiar with microformats. It’s not clear to me how microformats work, and how microformats and Dublin Core can interoperate on the web is clearly a very interesting question. When I look up the microformats.org web site, I understand that the difficulties addressing the encoding of Dublin Core as microformats in xhtml come from microformats themselves. They are a plurality of standards, not one homogeneous approach to doing a specific thing, like, say Dublin Core is. They are commonly (?) thought to be a web 2.0 technology, whatever web 2.0 means.

Now this is where the first paper enters. Didac Margaix Arnal applies his didactic genius in his paper “Concepts of web 2.0 and library 2.0: origin, definitions and challenges for today’s libraries”. Behind this boring sounding title hides a lively, well-written and interesting paper that introduces the issue. A rather fine bibliography complements the paper. If you are like me, you will be glad you read the paper. It helps tremendously to clarify ideas. I won’t tell you the end of the paper though, read it now yourself.

Thomas Krichel
krichel@openlib.org
http://openlib.org/home/krichel

College of Information and Computer Science
CW Post Campus of Long Island University
720 Northern Boulevard
Brookville NY 11548-1300, EUA
Faculty of Information Technology
Novosibirsk State University
2, Pirogova Street
630090 Novosibirsk, Rusia

Versión online de EPI

Existe una versión electrónica de El profesional de la información, de uso gratuito para todos los suscriptores de la revista, que permite acceder a través de internet a los textos completos y materiales gráficos publicados desde el año 2000.

Más información en:
http://www.elprofesionaledelainformacion.com/contenidos.html