uocpapers e-Journal on the Knowledge Society

http://uocpapers.uoc.edu

Iss. 7 | October 2008 | DOSSIER

"Web Science" Dossier

Universitat Oberta de Catalunya

uocpapers e-Journal on the Knowledge Society

http://uocpapers.uoc.edu

"Web Science" Dossier

Presentation

Julià Minguillón

Submission date: September 2008
Accepted in: September 2008
Published in: October 2008

The Internet is now a constant in all parts of life and at all levels: personal, academic and professional. In only a few years, the Internet has changed the way we relate, study, work, shop, spend our free time, etc. The web, the collection of resources accessible over the Internet, has become a gargantuan oracle where people search, create and share information on any subject, and where they communicate globally, breaking down the barriers of time and space. The possibility of being connected almost all the time from anywhere means that users have a perception of the web as something that is always available to cover any need. The rapid evolution of the technology available, both hardware and software, together with its relatively low cost, means that ever more users access the web and use it for more reasons.

Web science, which has become a hot topic recently, looks to explain the evolution of the web as a complex organism and an independent ecology, though multidisciplinary analysis of the range of knowledge areas that converge in the web. These are mainly technological (how does the web evolve technologically), organisational (what standards and specifications support this evolution) and social (what use do users make of this technology). There are also other areas involved in these changes, given that many economic, political and legal interests do not want to be left out of the evolution of the information society. Web science looks to provide a reference framework to ensure the proper analysis of all of these events from a range of perspectives and different levels of resolution, taking into account all the elements involved and the relationships established between them, with a clear multidisciplinary spirit to promote the participation of the widest possible selection of members of the scientific community.

This dossier presents what is meant by Web science through the following contributions:

- "Science, the Web and Web Science", by Daniel Riera, where the author provides an introduction to the concept of Web science, reviewing the first works on the subject that have appeared in recent years.
- "Trust on the Web: Some Web Science Research Challenges", by Kieron O'Hara and Wendy Hall, where the authors present the state of the art in Web science and the challenges set by multidisciplinary analysis of a range of quotidian fields.
- "E-learning from the Perspective of Web Science: Looking to the Future", by Julià Minguillón, where the author introduces e-learning as a clear example for the case study of the web, as an evolving field (distance education) where technology has played a part, as have methodological, organisational and other factors.

Julià Minguillón
Dossier Coordinator
Assistant Director of the IN3 (UOC)

Universitat Oberta de Catalunya

uocpapers e-Journal on the Knowledge Society

http://uocpapers.uoc.edu

article

"Web Science" dossier

Science, the Web and Web Science

Daniel Riera

Submission date: September 2008
Accepted in: September 2008
Published in: October 2008

Abstract

The evolution of the World Wide Web (or simply the web) since its birth in the 1980s has been staggering. The incorporation of new users, which has been exponential over time, and the constant contributions of new ways of organising data, communicating, sharing information, etc., have led to its gaining a practically unstoppable inertia in just two decades. The Web Science Research Institute, WSRI, was created in November of 2006 with the main aim of proposing a new discipline, Web Science, to observe the web and all that that surrounds it. It defends the need to analyse what is going on inside and outside the web and, thus, be able to propose improvements and corrections. This idea requires combining disciplines that have, to date, been very disperse, such as IT, psychology, law or economics. This has led to a new professional profile, the Web Scientist, and, likewise, new academic needs. This article details the work carried out by the WSRI in its first two years of existence, with regard to the new knowledge area of Web Science.

Keywords

Web Science, historical review, education, WSRI

Resum

L'evolució del World Wide Web (o simplement web) des del seu naixement a la dècada dels vuitanta del segle xx ha estat gegant. La incorporació de nous usuaris, exponencial en el temps, així com l'aportació constant de noves maneres d'organitzar les dades, comunicar-se, compartir informació, etc., han fet que el web hagi assolit en només dues dècades una inèrcia pràcticament imparable. El novembre de 2006 es crea la Web Science Research Initiative (WSRI) amb l'objectiu principal de proposar una nova disciplina, la «ciència del web» (Web Science), que observi la xarxa i tot allò que l'envolta. Defensa la necessitat d'analitzar què passa dins i fora de la xarxa i, d'aquesta manera, poder-hi proposar millores o correccions. Aquesta idea implica combinar disciplines tan llunyanes fins ara com la informàtica, la psicologia, el dret o l'economia. Tot això implica l'aparició d'un nou perfil professional, el científic del web, i, per tant, de noves necessitats acadèmiques. Aquest article mostra la feina feta per la WSRI en els primers dos anys d'existència amb relació a la nova àrea de coneixement: la ciència del web.

Paraules clau

Web Science, ressenya històrica, educació, WSRI

The inertia of the web

According to Sir Isaac Newton's first law of motion (1643-1727), "every body perseveres in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed". This first law goes by the name of the law of inertia[1] and attempts to explain the attitude of systems when faced with movement: if they are at rest, they will make an effort to stay at rest, and if they move, they will try to continue the movement. Therefore, depending on the object (or system), we can imagine, without going too deeply into the physics behind it, that this object will have more or less "resistance" to change. Consequently, if we have a football and a cannonball of the same size, both placed side by side, it would seem easier to make the former move than the latter. Equally, it will be easier to stop a lorry that is out of control at 1 kmh^{-1} than a motorbike, which is much smaller, at about 200 kmh^{-1}.

1. *Inertia*, in Latin, means leisure or laziness.

When Sir Timothy John Berners-Lee proposed the idea of connecting hypertext with TCPs and DNSs in 1989, he was making available to everyone (as the infrastructure already existed) what we now call "the web" (or the World Wide Web). In order to find an easy way of instantly sharing and updating information between scientists, a network was created that would enable contents that were far apart to be viewed as though they were on an individual's own computer. From this moment on, and with this little shove, a small snowball began to roll.

In next to no time (less than two decades), the web has gained inertia. First, the incorporation of the academic and, then, the business and private fields have expanded the size of the invention. In addition, increased features and the lowering of prices, linked to swift developments in hardware, and the incorporation of new, highly attractive ideas into the web, have enabled it to mature and become enriched very quickly.

If, initially, the web was viewed as the connection of a group of static (or almost static) contents, these contents have become increasingly dynamic. The web is now talked of as the Web 2.0[2] and appears to want to become 3.0 (Markoff, 2006). The current second version leans towards creativity, information sharing (not only among scientists) and collaboration through social networks, wikis or blogs, for example. Many of these components had been in existence since the web began, though they were not so widely known about or used. The main proposal of the future (third) version is to transform the use of the web's information into a semantic web (Berners-Lee *et al.*, 2001), with all the data globally interconnected and accessible, and with search engines that incorporate artificial intelligence. Likewise, the content will be primarily open, as will users' identities, and must be portable to any machine connected to the web.

The entity that the web has become, from the initial shove that it received from Tim Berners-Lee some 20 years ago, is immense. Large-scale expansion began in 1995, when it already had 16 million users. By 2001, there were 400 million and 2005 saw it reach the first billion-user mark. It is expected that there will be more than two billion[3] by 2010.

These two types of exponential growth (qualitative and quantitative) make the accumulated inertia seem uncontrollable. A good test of this is that, in terms of teaching, what a few years ago could just be summarised in a couple of lessons in an IT degree subject, is now a volume that could constitute a complete degree all by itself.

Given all this, on November 2 2006, four researchers from MIT and the University of Southampton founded the Web Science Research Initiative (WSRI). In their presentation, they argued the need for a new interdisciplinary field – which would embrace mathematics, IT, psychology, econòmics, law, etc. – to analyse what the web is and what happens on it and around it. A field showing the vulnerabilities, the interactions that need improving, the implications that microscopic decisions have on global design, etc. A science that works on both the technical and social aspects. They also suggested that this new area be led by specialists, with the proviso that they be skilled in this wide range of subjects. This requires the appearance of teaching devoted specifically to Web Science.

The Web Science Research Initiative

There is a great proliferation of international meetings linked in one way or another with the web. Despite this, it was not until the creation of the WSRI that the web began to be treated in a global and multidisciplinary way. From the start, the WSRI has tried to broadcast its purpose through publications and presentations at international congresses and conferences. It has tried to bring together experts from different fields so that they can all start sharing the idea of Web Science.

To achieve this, The Web Science Workshop was held in 2005 at the London offices of the British Computer Society[4]. The Emerging Science of the Web[5] was attended by 21 participants. The central topic of the encounter was the exploration of the critical challenges of the future of the World Wide Web (Hendler, 2005). This was the first global formal discussion on Web Science.

Subsequent workshops were held exclusively on this subject at renowned conferences, such as the Seventeenth[6] and Eighteenth[7] International World Wide Web Conferences. During these workshops, participants were asked to look at the future of the web and discuss how this web may be observed, analysed and influenced. Similarly, Hypertext 2008 (HT2008)[8] included the Web Science workshop: Collaboration and Collective Intelligence,[9] which primarily discussed new types of collaboration on the network, its structure and user behaviour.

Many other talks and presentations have been held between these workshops, which has helped the WSRI and its members

2. Not everyone likes this name. For example, Tim Berners-Lee does not think it appropriate. In any case, it seems to be widely accepted generally by the community.
3. See: <http://www.internetworldstats.com/stats.htm>.
4. See: <http://www.bcs.org/>.
5. See: <http://www.cs.umd.edu/~hendler/2005/WebScienceWorkshop.html>.
6. See: <http://www2008.org/>.
7. See: <http://www2009.org/>.
8. See: <http://www.sigweb.org/ht08/>.
9. See: <http://users.ecs.soton.ac.uk/dem/workshops/webscience08/>.

to start spreading their ideas (Weitzner *et al.*, 2007; Hall, 2008; Hendler *et al.*, 2008). In addition, some of the founding fathers of the WSRI have been interviewed or have written for general interest newspapers (Cellan-Jones, 2008; Lohr, 2008) or for scientific publications (Schneiderman, 2007; Hendler *et al.*, 2008b). At the Web Science Curriculum Workshop in September 2008, the debate concentrated on the objectives of the initiative that are associated with the teaching of Web Science, and the First International Web Science Conference, which is to be held in Athens in March 2009, has been planned for the near future.

A quick search on Google for "Web Science" brings up around three and a half million pages on the subject. Apart from the congresses, articles and interviews that have already been mentioned, these pages show blog entries in a number of languages (Saravanan, 2007; Peña, 2007; Ragget, 2008), news and articles in online journals and newspapers (Biever, 2006; LaMonica, 2006), and some definitions in wikis. Most of these articles and blog entries are from November 2006 and only relate to the initial WSRI announcement.

One of the last (and first) texts published is the provocative book *The Web's Awake* (Tetlow, 2008), which presents the web as a new way of life that we can no longer control. It provides a description of new web characteristics and types of behaviour that we have not created, but which have emerged of their own accord.

Interestingly, we should point out that the Wikipedia[10] definition of Web Science is rather poor. In line with the spirit of the web 2.0, it is users who should complete this content.

A Framework for Web Science

Given that this is a young field, little has been written about Web Science. Of the few references that we have, *A Framework for Web Science* (Berners-Lee *et al.*, 2006) is possibly the most complete. It is the book[11] with which the WSRI has tried to establish the foundations of the idea of Web Science. It deals with a number of core issues in the development of the web without going into excessively technical details. Given the quantity and range of knowledge in the web environment, it does not appear easy to bring it together under a single, homogeneous heading. For this reason, the aim of the book is not to be a collection of this technical knowledge, but rather a more informative introduction to the main topics that would constitute Web Science.

Consequently, after a brief introduction, the science and the engineering behind it, analysis of the web, social aspects involved and finally the legal, governance, security and standards issues are presented. A great effort is made to establish the connection between very wide-ranging subjects which so far have been analysed in isolation, or, if not, at least independently. Although these subjects are not fully linked, we are beginning to see a structure of the new science proposed.

Some of the aspects emphasised in the book are related to web organisation, from both the social and content, or infrastructures and resources points of view. This underlines the importance of well-designed web architecture. There is also a commitment to organisation and intelligent access to content (or information) through the use of labels, the Semantic Web, etc. It describes the importance of using mathematical models to analyse the web, which has to lead to continuous improvement. In addition, the implication of the social, cognitive and moral contexts of the web on its engineering decisions are also noted. Finally, it asks what type of regulation should be applied to the web and its users. Primarily, this includes areas related to user security (e.g. privacy and identity), improper behaviour, the importance of everyone accepting standards and the search for the web's governing structure.

The main shortcoming of the text is that the heterogeneity of the topics that it deals with makes a homogeneous study of the discipline itself difficult. At times, the text seems to be simply a review of a collection of well chosen articles. It would possibly have been more interesting if greater effort had been made to coordinate the different texts. Despite this, it is a valuable work that sets out the bases of Web Science. It covers a range of subjects that are fairly representative of what the web now is and the directions it has taken. However, it does seem strange that despite covering so many subjects there is not the slightest mention of the web's influence on teaching, or concepts such as e-learning or distance education when this is one area where the development of the web has had the most influence.

Academic contributions

One of the objectives of the WSRI, if not in the short term then in the medium term, is the joint design between the MIT and the University of Southampton of a Web Science degree. To date, this degree has not been offered and there is no indication as to when it will be up and running or the progress made so far.

An initiative closely associated with Web Science is iSchools (or Schools of Information),[12] presented in September 2005 at Pennsylvania State University. This proposal differs from Web Science in that it focuses more on relations between information,

10. See: <http://en.wikipedia.org/wiki/Web_science>.
11. It is in fact not strictly a book, but the first edition of the journal on Web Science: *Foundations and Trends in Web Science*.
12. See: <http://www.ischools.org/oc/>.

technology and people, irrespective of the platform (which does not necessarily have to be the web). The Oxford Internet Institute[13] also offers training on the web, though this is aimed primarily at Masters and Doctoral level. Likewise, there is the thread concept as proposed by Georgia Tech,[14] which offers students the possibility to define a syllabus, which, although some way away, resembles to some extent a Web Science degree. In any event, these threads are more akin to what is now seen as a specialisation within a degree, and not a degree in itself.

Finally, the proposed future IT Degree from the UOC includes an optional subject devoted to Web Science. Work is also underway to prepare postgraduate teaching in this field.

Conclusions

On the basis of the WSRI proposal and the most easily accessible information to date, it seems that Web Science is still at an early stage of development. Its members continue to strive to make the proposal more widely known and give the impression that the core discussion subject is still somewhat disperse. Clearly, it is not easy to give a precise definition of what comes under this new area, what is left out, and what interactions there are among the sub-areas.

All this does not make designing the new associated qualifications easy. At present, MIT and the University of Southampton are supposedly preparing the first Web Science degree, while the remaining universities await it. The syllabus of this new degree has to be widely accepted by a significant number of university departments, which complicates the design.

Despite this, the proposal is only two years old and significant efforts seem to be underway to bring it into being and provide a solid definition of the new field. The importance that the web needs to take on in twenty-first-century society could be compared to the significance of penicillin, the printing press, the steam engine or electricity in previous centuries. Web Science is essential to ensure the creation of professionals who are capable of understanding and influencing the web and that which surrounds it.

Reference

BERNERS-LEE, T.; HAENDLER, J.; LASSILA, O. (2001). "The Semantic Web" [online article]. *Scientific American*. [Accessed: 20/08/2008]. <http://www.sciam.com/article.cfm?id=the-semantic-web>

BERNERS-LEE, T.; HALL, W.; HENDLER, J. A. [*et al.*] (2006). "A Framework for Web Science". *Foundations and Trends in Web Science*, vol. 1, no. 1.

BIEVER, C (2006). "Web pioneers call for new 'web science' discipline" [online article]. *New Scientist Tech.* [Accessed: 20/08/2008]. <http://technology.newscientist.com/article/dn10438-web-pioneers-call-for-new-web-science-discipline.html>

CELLAN-JONES, R. (2008). Article and interview with Dr Wendy Hall [online]. BBC [Accessed: 20/08/2008] <http://www.bbc.co.uk/blogs/technology/2008/03/what_is_web_science.html>

HALL, W. (moderator) (2008). *Does the Semantic Web Need Web Science?* [online] Debate panel at the 5th European Semantic Web Conference, Tenerife. [Accessed: 20/08/2008]. <http://videolectures.net/eswc08_hall_dsw/>

HENDLER, J. (2005). *Web Science Workshop. The Emerging Science of the Web* executive summary [online]. [Accessed: 20/08/2008]. <http://www.cs.umd.edu/~hendler/2005/WebSciReport.pdf>

HENDLER, J.; BERNERS-LEE, T.; HALL, W. [et al.] (2008). Presentation of the *Tetherless World Constellation: The Future of the World Wide Web* [online]. [Accessed: 20/08/2008]. <http://www.rpi.edu/news/events/tw/index.html>

HENDLER, J.; BERNERS-LEE, T.; HALL, W. [*et al.*] (2008b). "Web Science: An interdisciplinary approach to understanding the World Wide Web". *Communications of the ACM*, 51, 7 (pp. 60-69).

LaMONICA, M. (2006). "Berners-Lee, universities launch 'Web Science' initiative" [online article]. *c/net News.com*. [Accessed: 20/08/2008] <http://news.cnet.com/Berners-Lee,-universities-launch-Web-science-initiative/2100-1025_3-6132016.html?hhTest=1 , 2 November 2006>

LOHR, S. (2006). "Group of University Researchers to Make Web Science a Field of Study" [online article]. *The New York Times*. [Accessed: 20/08/2008]. <http://www.nytimes.com/2006/11/02/technology/02compute.html>

MARKOFF, J. (2006). "Entrepreneurs See a Web Guided by Common Sense" [online article]. *New York Times*. [Accessed: 20/08/2008]. <http://www.nytimes.com/2006/11/12/business/12web.html>

PEÑA, I. (2007). *Last Thoughts About Web Science And Academic Blogging Or Why Did Not Academia Came Up With Wikipedia. And Some Acknowledgments Too* [online article (blog)]. [Accessed: 20/08/2008].

13. See: <http://www.oii.ox.ac.uk/>.
14. See: <http://www.gatech.edu/>.

<http://ictlogy.net/review/20070729-oii-sdp-2007-epilogue-last-thoughts-about-web-sciences-and-academic-blogging-or-why-did-not-academia-came-up-with-wikipedia-and-some-acknowledgments-too/>

RAGGET, D. (2008). *Web Science and link rot, a way forward* [online article (blog)]. [Accessed: 20/08/2008]. <http://people.w3.org/~dsr/blog/?p=7>

SARAVANAN (2007). *Web Science: Web is a network of people* [online article (blog)]. [Accessed: 20/08/2008]. <http://saran945.com/blog/web-science-web-is-a-network-of-people/>

SHNEIDERMAN, B. (2007). "Web science: A provocative invitation to computer science". *Communications of the ACM*, 50, 6 (pp. 25-27).

TETLOW, P. (2008). *The Web's Awake: An Introduction to the Field of Web Science and the Concept of Web Life.* Piscataway: IEEE Press.

WEITZNER, D.; ALONSO, J. M. (2007). Summary of the *Toward More Transparent Government Workshop on eGovernment and the Web Workshop* [online]. [Accessed: 20/08/2008]. <http://www.w3.org/2007/06/eGov-dc/summary>

Recommended citation

RIERA, Daniel (2008). "Science, the Web and Web Science". In: Julià MINGUILLÓN (coord.). "Web Science" [online dossier]. *UOC Papers*. Iss. 7. UOC [Accessed: dd/mm/yy].
<http//www.uoc.edu/uocpapers/7/dt/eng/riera.pdf>
ISSN 1885-1541

Daniel Riera
Director; IT Engineering Programme; IT, Multimedia and Telecommunications Department; UOC.
drierat@uoc.edu

PhD in IT Engineering from the Autonomous University of Barcelona (UAB) in 2006, Masters in Advanced Process Automation Techniques in 2002 from the UAB and degree in IT Engineering, also from the UAB, in 1998.

As a Ministry scholarship holder with a Research Personnel Training grant, he spent a year at Imperial College London, as a visiting scholar, working in the ICParc planning and resource control department. It was here that he received training in constraint programming and optimisation, and collaborated on projects for the RAC and Railtrack.

He has taught at the UAB from September 1998 to August 2005 and since September 2007. In September 2005, he became a lecturer in the UOC's IT, Multimedia and Telecommunications Department. Since September 2006, he has been working as the Academic Director of the Postgraduate course in Bio-computing currently offered by the UOC. His teaching work focuses primarily on programming and bio-computing. In January 2007, he was named director of the department's IT Engineering Programme.

His main areas of research include discrete systems modelling using Petri nets and optimisation with constraint programming techniques. He is the author of a number of articles and papers for national and international journals and conferences. He has participated in research projects and agreements as a member of the research team. He has worked as a researcher with the LOGISIM, the Centre for Logistics Systems Simulation and Optimisation, part of the CIDEM's Network of Technological Innovation Support Centres. He is currently a member of the UOC's software engineering research group (GRES-UOC), where he is researching verification of UML/OCL models using constraint programming.

Universitat Oberta de Catalunya

uocpapers e-Journal on the Knowledge Society

http://uocpapers.uoc.edu

article

"Web Science" Dossier

Trust on the Web: Some Web Science Research Challenges

Kieron O'Hara and Wendy Hall

Submission date: September 2008
Accepted in: September 2008
Published in: October 2008

Abtract

Web Science is the interdisciplinary study of the World Wide Web as a first-order object in order to understand its relationship with the wider societies in which it is embedded, and in order to facilitate its future engineering as a beneficial object. In this paper, research issues and challenges relating to the vital topic of trust are reviewed, showing how the Web Science agenda requires trust to be addressed and how addressing the challenges requires a range of disciplinary skills applied in an integrated manner.

Keywords

Web Science, trust, privacy

Resum

La ciència del web és l'estudi interdisciplinari del World Wide Web com a objecte de primer ordre per a entendre la seva relació amb les societats més grans en què és integrat i per a facilitar la seva enginyeria futura com a objecte beneficiós. En aquest article es presenten els problemes i els reptes de la recerca sobre el tema vital de la confiança i es demostra que l'agenda de la ciència del web exigeix que es tracti la qüestió de la confiança i que per a enfrontar-se a aquests reptes cal un ventall de destreses disciplinàries aplicades d'una manera integrada.

Paraules clau

ciència del web, confiança, confidencialitat

Introduction: Web Science

The Web is one of the most ubiquitous and transformative technologies humankind has ever produced, and in a remarkably short space of time has been embedded into a very large number of social activities, ranging from e-science to e-commerce, from e-government to entertainment, from citizen journalism to cybercrime. Yet we still remain in remarkable ignorance of the trajectory of the Web's development (and, conversely, of what risks the Web faces given its major exposure to the world); now-mainstream activities, such as blogging, file-sharing or social networking were unheard-of just a few years ago, and their appearance and sudden blossoming took most scientists and pundits by surprise.

There are doubtless many reasons for this, not least being the rapid inflation of such activities from niche practices to mainstream behaviour. But one of the most important is that we are lacking the conceptual tools and focused effort required to understand the Web. Of course it is a piece of computer engineering, but it is not simply the sum of TCP/IP, HTML, HTTP, PageRank, Ajax, URIs and whatever else. It is also created, written, linked and read by hundreds of millions of people.

Hence, the Web is beyond the purview of any individual discipline, even computer science. Google's search algorithm PageRank is an impressive piece of work, but understanding the algorithm does not tell you about its place in the Web. For that you would need to understand the function of search, the complex real-world environment plagued by bad behaviour, such as Google-spoofing, the economics of Google's click-based advertising business model, the engineering of Google's indexing and caching methods and so on. Web Science transcends faculties, requiring theoretical science, empirical science, engineering, social science and humanities. In 2006, a group of computer scientists launched the Web Science Research Initiative (WSRI)[1]

1. Visit: <http://webscience.org/>.

to help create the interdisciplinary study area that would support our understanding of the Web. As the founding directors of WSRI put it:

> Web science is about more than modeling the current Web. It is about engineering new infrastructure protocols and understanding the society that uses them, and it is about the creation of beneficial new systems. It has its own ethos: decentralization to avoid social and technical bottlenecks, openness to the reuse of information in unexpected ways and fairness. It uses powerful scientific and mathematical techniques from many disciplines to consider at once microscopic Web properties, macroscopic Web phenomena, and the relationships between them. Web science is about making powerful new tools for humanity, and doing it with our eyes open. (Berners-Lee *et al.*, 2006)

In this paper we wish to present an example of a particular problem whose solution(s) require the full breadth of scale and disciplinary experience to which Berners-Lee *et al.* allude: the problem of trust on the Web.

There are a number of reasons why trust is an issue online: the decentralised environment, the lack of supporting contextual factors, the artificiality of many agents, the fluidity of identity, the highly heterogeneous user base. But most importantly trust is essential. Should you trust the content you download, given the lack of a central moderator? If you use automated services, should you rely on their effectiveness, and should you give them access to sensitive information? How do you know that people are representing themselves accurately in social networks? Any Web user must learn to place trust in content, in services and in people wisely and safely. Jennifer Golbeck (2006) lists three major challenges to applying trust online. *Trust management* is the process of determining who has access to which information or resources. *Trust computation* is the method of deriving a level of trust in a resource on the basis of the available data. *Integrating trust into applications* involves building applications that can function by placing trust accurately enough for its purposes.

In the remainder of this paper, we will survey research challenges relating to online trust from the point of view of Web Science, showing that a road map for Web Science research into the problem of trust will among other things have to cross varying scales and disciplines. We begin philosophically and psychologically, examining the features of trust and its management that would ideally be replicated online. Next, we switch to sociology, and examine some of the attitudes to trust that prevail among Web users. The next section takes the position of politics and security, looking at the relationship between trust and the related concept of privacy. Finally, we bring in technology, examining developments to promote trust on the Semantic Web. In the conclusion, we discuss how "the problem of trust" has, from a point of view consistent with a traditional division of disciplines, decomposed into several research challenges. We take this as *de facto* evidence that the inclusive approach of Web Science is essential.

Trust and trustworthiness

Without trust the Web wouldn't function; exchanges of resources or information require all sorts of risk-taking. But what do we get in return for the risk? One clear function of trust is as a method of complexity reduction (Luhmann, 1980). Trust enables us to ask or pay others to act on our behalf. If we trust them, we do not have to carry out the subtask ourselves, and neither do we have to monitor or micromanage their performance. It enables us also to ask advice or receive instruction; if we do not trust our advisors, we will have to acquire their expertise before we can act confidently. So, although there is a considerable body of evidence that trust has a large moral component (Uslaner, 2002), trust performs the important social function of increasing the efficiency of social interaction. Online, the moral dimension to online trust is less well established and may not be very relevant at all (social networking may be changing this). Online trust is generally reduced to an evidence-based cost/benefit/risk analysis of expectations of whether performance will live up to our expectations (see the survey of methods in Golbeck, 2006).

Connecting trust and trustworthiness

Nevertheless, although trust is a good thing and commentators have argued that high-trust societies have advantages over low-trust societies (Fukuyama, 1995), we should note that increasing trust is not a solution to all social problems. There is an important distinction that is often blurred in discussion between trust and trustworthiness (*cf.* Hardin, 1996). Trust is an attitude of one agent, X, to another, Y. Trustworthiness is a property of an agent. These are usually relativised to a task – one person trusts another *to do something particular* (we will ignore this caveat in the rest of our discussion). So, X trusts Y just in case X *believes* Y is trustworthy.

X trusting Y will certainly reduce the complexity of X's life, but if X's belief is actually false, then the cost to X in resources risked in the transaction will be correspondingly high. What is just as important to X is that her belief about Y is actually true (*viz.*, that Y actually is trustworthy). Conversely, Y's trustworthiness does him no good as long as no-one believes he is. There is an opportunity cost to being trustworthy without trust.

The incentives are skewed in a difficult way. X benefits from Y's trustworthiness, but controls only her trust; Y benefits from X's trust, but controls only his trustworthiness. What is essential – what a functioning society does – is to link trust and

trustworthiness so that ideally all and only trustworthy people are trusted. The challenge, then, is not how to increase trust, but rather how to create a causal link between trustworthiness and trust. This presents us with the first set of Web Science research challenges.

- How should we maintain a causal link between trust and trustworthiness? What incentives and economic models are available to promote trust and trustworthiness together?

Local and global trust

Trust can be decomposed into *local* and *global* trust (O'Hara, 2004), depending on the evidence one uses for supporting one's beliefs about the trustworthiness of others. Local trust relies on personal acquaintance mediated through high-bandwidth interactions. We take note of varying signals given out by agents – facial expressions, clothing, language and so on. These signals provide the causal link that connects trust and trustworthiness. The downside of the system is that once a signalling convention (wearing a suit, say) has been adopted and codified in a society, then it can be faked by untrustworthy people. Hence trustworthiness signalling systems need to be updated constantly in an arms race with untrustworthy counterfeiters.

Global trust involves outsourcing our trust decisions to trusted institutions, so X would trust Y on the basis of a certificate from an institution that Y was trustworthy. Such institutions do not "solve" the problem of trust – they merely shift it, because X must decide whether she trusts the certifying institution, not Y. But the institutions do have the important effect of globalising trust, because X can take informed trust decisions about people she has never met. An institution has all sorts of economies of scale that enable it to perform more thorough investigations of its subjects, and so it can establish a very strong connection between trust and trustworthiness. However, the downside is systemic risk – one mistaken certification could lead to all the institution's output being ignored and its clients withdrawing their trust from everyone it has certified.

The variables used for gathering evidence will affect the type and therefore the scope of trust. For instance, Huynh *et al.* (2006) present a composite model where four types of mechanism are combined to produce a decision to trust or not to trust in an online agent context.

- Interaction trust. Based on past experience of direct interactions.
- Role-based trust. Defined by role-based relations between the parties.
- Witness reputation. Based on reports from witnesses of past behaviour of the agent.
- Certified reputation. Based on third-party references provided by the agent.

Broadly, local trust is provided by the first of these, the actual experiences of the trusting party in his/her history, and current experiences in looking at eg, a website. The other mechanisms supply global trust (eg, the roles the two parties play such as citizen/government, customer/retailer, customer/bank, ISP/customer allow the parties to go beyond personal experience to place their trust).

The decentralised nature of the Web, plus its rapid changes, mean that there are few trusted institutions online. Many institutions trade on their offline reputations (for instance, many banks, universities and government institutions), while very few have developed a reputation entirely online (PayPal would be one example). So despite the global nature of the Web online trust is often local, in that it relies on a person's personal experience in dealing with a website. A user interacts with a website and assesses the signals given out by that website him- or herself.

There are two obvious problems with this. First, online the user is deprived of the complexity of signal available in the offline world (which include quite unconscious signals of (un)trustworthiness such as a shaking hand or unconfident expression). Online, the signals are basically the visual ones described by the HTML source file of the page, augmented possibly by the roles played by the parties in the transaction (*cf.* role-based trust), which in general is not a secure source of trust on its own. And second, the designer of the website is in total control of the signals that it gives out; the user has little or no opportunity to engage the website in "conversation", to see how it "performs", to "size it up", as we do offline when we are judging people. When trust is local, based on personal acquaintance, the dice are loaded in favour of the website in these two ways. This presents us with a second set of research challenges.

- How should trust be represented, maintained and repaired on the Web? What variables are important? Will these change as we move from human to artificial agents? What sort of institutions and methods will globalise online trust?

Bootstrapping

Sometimes the trust/trustworthiness link needs to be started on the basis of little evidence – the so-called *bootstrapping problem*. Y sets up in business – should X trust him? She cannot without evidence that Y conducts his business in a trustworthy fashion and when Y has just started there is no such evidence. But unless someone trusts Y to work for them, there will never be such evidence.

Offline, we have a number of strategies – many exploiting the moral nature of trust. There has been a lot of debate between the Weberian idea that trustworthiness causes trust, and the Durkheimian position that trusting people inclines them to behave in a trustworthy manner. In fact, we muddle by with a combination of the two, breaking into the circle and bootstrapping the relationship. We rely on moral notions such as duty and inclusion of people into our moral community.

But online, bootstrapping is a problem, partly because of the relative unimportance of the moral dimension (with the caveat that social networking may be bringing a more Durkheimian model of trust with it). Most models of online trust are evidence-based and it is evidence that is lacking in the first place. In terms of the model of Hyunh *et al., ex hypothesi* there is (a) no past experience of direct interactions, (b) a lack of experience in any meaningful role, (c) no witness reports, and (d) no evidence-based third-party references. But trust has to begin, somehow. This gives us another set of research challenges.

- How can effective and accurate trust be bootstrapped?

Web users

Bootstrapping is also important when we consider the Web as a social system, rather than a collection of linked hypertext documents and data. Trust in Internet transactions is higher, unsurprisingly, with Internet users than non-users (Dutton *et al.*, 2007); the growth of the Web does depend on those users achieving a level of familiarity in order to reduce uncertainty and increase their confidence. Even ex-users of the Internet trust it more than non-users.

People are remarkably accepting of bad experiences online, such as spam or obscene email, at least up to a point. Such "anti-social behaviour" is assumed to reduce levels of trust, so this might seem a surprising result given the prevalence of spam etc., as reported in the media and academic work. One explanation is that those providing such reports tend to be heavy users of the Internet, and therefore much more likely to have bad experiences; general patterns of use must also include a large number of people who use the Internet relatively rarely, but whose voices are seldom heard in discussions about the Internet or the Web (Dutton *et al.*, 2007).

As one would perhaps expect from the previous section, people's online trust mirrors the experience of those placing local trust, ie, unmediated and based on personal acquaintance (in the online case, with a website). For instance, one review of online trust discerned three *perceptual* factors that were particularly relevant. *Perception of credibility* is to do with honesty, expertise, predictability and reputation. *Ease of use* relates to the simplicity and design of the website. *Risk* is the perceived likelihood of an undesirable outcome (Corritore *et al.*, 2003). Risk is of course a pervasive issue with trust, but the other two factors are strongly connected to the gathering and evaluation of signals of trustworthiness. Credibility signals are designed to display the trustee's expertise, but ease-of-use signals, which include having a well-designed site, avoiding such pitfalls as bad spelling and dangling links, are strictly unconnected with credibility and easy to fake, yet are still important signals. This confirms the findings of an earlier study which found six major features that encouraged trust in e-commerce sites – the site's brand, seals of approval, ease of navigation, a fulfilling ordering experience, the site's presentation and the technologies used to create the website – which again are strongly connected with the signalling systems characteristic of local trust (Cheskin Research, 1999).

However, Web users are not particularly efficient at picking up the right signals that provide the causal connection between trust and trustworthiness. Dhamija *et al.* (2006) investigated the reasons why bogus sites work and discovered that existing anti-phishing browser cues – the "signals" which users are supposed to pick up and which connect trust and trustworthiness – are ineffective. A participant group in that experiment made mistakes 40% of the time (even though primed to look out for phishing sites), and surprisingly neither age, gender nor computing experience were significant variables. The study showed that people are unaware of the sorts of signalling systems that have been developed to ensure trustworthiness (eg, the padlock symbol to show that the page was delivered securely by SSL), or of the typical strategies of counterfeiters (eg, using images to mask underlying text, or placing an SSL-padlock in the body of a webpage). Furthermore, users often fail to notice the *lack* of expected signals of trustworthiness. Attention to the needs of actual Web users leads to a further set of Web Science research challenges.

- How can secure systems be made usable and effective for consumers, given the bounded knowledge and rationality of Web users?

Privacy

One of the biggest obstacles to trust on the Web is the threat that digital information provides to the user's privacy (O'Hara and Shadbolt, 2008). Thanks to the Internet and the Web, information is very easily copied and transferred. These technologies were founded on a liberal ideology of free-flowing information, conceived in the context of fast-moving scientific research, where access to data was limited and the publishing cycle was slowing down the research cycle. The Web provided a means for

information-sharing that has clearly boosted, indeed transformed, research.

However, outside the realm of science, information, particularly personal data and intellectual property, has a value which may be threatened by copy and transfer. The trade-off between the benefits and costs of freedom of information is an acute problem on the Web (though not unique to it). It is clear that problems with online privacy are perceived as the major risks by users (Dutton *et al.*, 2007). Furthermore, as in general wealthier and better-educated people are more likely to be experienced online users, there is an unequal distribution of privacy whereby privacy-aware individuals have the means to protect privacy (at some cost), while people with a strong Web presence who remain unaware of the risks or unable to understand or afford protections are potential targets (O'Hara *et al.*, 2006).

This is a classic Web Science problem, at the interface between group behaviour, individual perceptions, the politics of information and Web engineering. Furthermore, many issues pertaining to security pop up here, too, such as the need to retain usability and to integrate solutions into standard workflow and Web use. One multi-layered attempt to address this problem is the *Policy-Aware Web* (PAW, Weitzner *et al.*, 2005), which suggests a mechanism whereby, if a browser requests information (for example, images to be displayed within a Web page), it receives a modified HTML 401 error which includes the URI of the site's privacy policy (a statement of restrictions of access to the information). The browser would then construct a proof, based on its own credentials, that the policy was satisfied. The site would then check the proof and release the information if the proof was valid.

Such an approach has a lot of promise, but there are still tensions within society as a whole between privacy and transparency. New types of interaction, including social networking, lifelogging and the maintenance of virtual identities will demand new assessments of the risks to privacy (Bailey *et al.*, 2007; O'Hara *et al.*, 2008) and consequent re-examination of the law (Allen, 2008). It is in this space that Web Science faces some of its most daunting challenges.

- What privacy issues arise in a Web environment of increasingly sophisticated information sharing? Can traditional forms of regulatory regime cope? To what extent are the service providers going to become the legal gatekeepers for public authorities in terms of delivering their public policy objectives, eg, Web policing for what is judged to be "illegal and harmful content"?

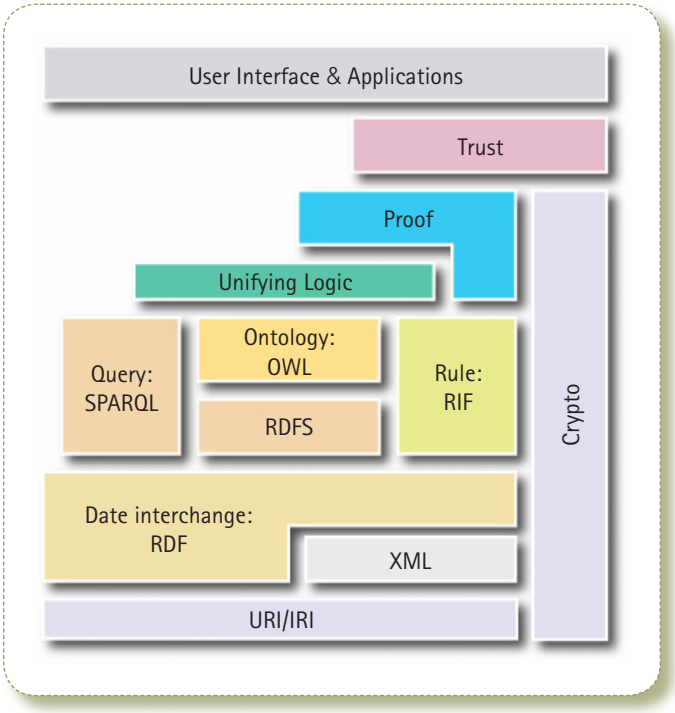
Privacy, trust and security issues loom large in a number of specific areas of Web activity, for instance e-health. The use and ownership of information set up another set of research challenges.

- What legal regimes are appropriate where users are heterogeneous and often inexperienced? For example, in e-health scenarios, where it is important that professionals have effective and timely access to information, who should own information about patients, where patients have an interest in privacy but many others have an interest in the information?

The Semantic Web, provenance and social networks

The use of the Web's own technology (URIs, theorem proving, error handling) to address privacy with the PAW points to the possibility of doing the same thing for trust. This is indeed essential, as projected developments to the Web, particularly including the Semantic Web (Shadbolt *et al.*, 2006) are premised on the inclusion of a technological layer to address the issue of trust. If we look at a layered view of the various formalisms and protocols which are currently seen as making up the Semantic Web (SW), trust has a vital place at the top of the edifice (Figure 1).

Figure 1: The Layered View of the Semantic Web



Work on the trust layer is at an early stage and is currently very fragmented, partly because of a range of opinion on what is likely to establish trust in the data which the SW is intended

to make available (Golbeck, 2006). It is certainly the case that without a trust layer, the amount of data which will be made available may well be less than the SW's designers hope – another example of the interaction between social pressures and technological change.

There are a number of approaches under research, including the proof-based idea implicit in the PAW. Another aspect is the creation of metadata about the provenance of an information resource, detailing where it came from and what methods were used to generate it; this metadata can be used to inform trust decisions. Moreau *et al.* (2008) describe a method of metadata provision that is crucially sensitive to the lifecycle of provenance, using an open data model for documentation of a resource which serves user queries over a representation of what processes were involved in generating that resource. The key point is that this idea is consistent with a number of computing trends towards open applications, composed dynamically which derive results on the fly.

A third method exploits a quirk about trust, which, although not transitive (if A trusts B, and B trusts C, it does not follow that A trusts C, or even that A should rationally trust C), has transitive qualities (A's trust of B might lead A to give B's opinion of C extra weight). So, for example, Richardson *et al.* (2003) use social networks with trust to calculate the belief a user might have in a statement by finding paths through a social network from the user to a node that represents the opinion in question. Trust values along the paths are concatenated and aggregated to provide a global trust value.

One influential SW project is Friend-of-a-Friend (FOAF), a framework for representing information about people and their social connections, enabling the iterative creation of a social network via connective predicates such as "A knows B" (Brickley *et al.*, 2007). Such a network can make it possible to make judgements about someone you don't know, purely by seeing their place in a network. For instance, one could decide to allow access to your information to people within two steps of you in the network. Or you could use a weighting algorithm to determine how trustworthy someone is likely to be, given that he or she is known in some role by certain people you trust. Golbeck *et al.* have added a trust module to FOAF allowing people to rate how much they trust each other, either with respect to a particular topic, or in general (Golbeck *et al.*, 2003). FOAF is beginning to be applied not just to the Web as a whole, but to the important sub-world of social networks (Golbeck *et al.*, 2008; O'Hara *et al.*, 2008).

The fact that there is a wide variance in methods of computing trust and a similar range of contexts in which that has to be done, means that a single means of dealing with the problem is unlikely – and consequently that there will not be a purely technological solution, and that the interdisciplinary range of Web Science will need to be leveraged. For instance, analysis of the costs and benefits of different strategies in the manner of the framework (O'Hara *et al.*, 2004) will be required; this framework does not draw conclusions but rather details the costs of various strategies of applying trust and also subdivides the costs into operational costs, opportunity costs, risk, costs of betrayals and service payments.

The question of bootstrapping trust also has to be taken into account – how should a new entrant "break into" a social network? Supplying provenance metadata has to be part of the answer, as does a policy-based approach. And one advantage of the social networking approach is that once one has established social relations with one or two of the network members, one does have some sort of network presence, however minimal, which provides an opportunity for more transactions potentially leading to more trust. O'Hara *et al.* (2004) also discuss which strategies for placing trust can help circumvent the bootstrapping problem. In general, the more optimistic the strategy, the better for bootstrapping. Centralised strategies also can work, but not only are they hard to scale, but they work against the Web's decentralised ethos.

Hence the technology presents us with more research challenges.

- What languages and ontologies are appropriate for expressing the requirements for online trust? At the moment, the work on trust on the Semantic Web is relatively sparse and unfocused – how should it be focused?
- How can accountability and transparency be engineered into information use? Given the ability to describe information policies, how can they be enforced? How can we ensure the quality of provenance metadata?
- How should the trust layer of the Semantic Web be integrated with the layers lower down to create a seamless interaction for the user?

What is fascinating is that this set of research challenges brings us right back to the start, and our first set of challenges about establishing and maintaining the causal link between trust and trustworthiness.

Conclusion

Trust cannot be engineered, but mechanisms can be put into place that aid standard mechanisms to create trust. As noted above, the Web has aspects which put trust at risk; it is by going with the grain of society that the Web can promote trust of itself and also support the globalisation of trust in wider society. To understand the grain of society, it is essential to understand the

interactions of online trust at a number of levels ranging from the micro (the protocols governing the transfer of information) to the macro (the social effects of information being passed around) and all stages in between (for example, the individual psychology of online trust).

There are many specific research challenges with which Web Scientists must wrestle, many of which we have highlighted and bulleted in this paper. Any or all of these would make fascinating research projects or PhD topics – but none of them can be properly approached from the perspective of a single discipline, even computer science. Our aim has been to indicate the broad range of disciplines required to understand the problem; it is doubtless not an exhaustive list, but in its breadth it is perhaps a very strong argument for the importance of Web Science for the future not only of the Web, but our Web-enabled society.

Reference

ALLEN, A.L. (2008). "Dredging up the past: lifelogging, memory and surveillance". *University of Chicago Law Review*. Vol. 75, pp. 47-74.

BAILEY, J.; KERR, I. (2007). "Seizing control? The experience capture experiments of Ringley & Mann". *Ethics and Information Technology*. Vol. 9, pp.129 -139.

BERNERS-LEE, T.; HALL, W.; HENDLER, J. [*et al.*] (2006). "Creating a science of the Web". *Science*. Vol. 313, pp. 769-771.

BRICKLEY, D.; MILLER, L. (2007). *FOAF Vocabulary Specification 0.91*. [Date accessed: 11/07/08].
<http://xmlns.com/foaf/spec/>

CHESKIN RESEARCH AND STUDIO ARCHETYPE/SAPIENT (1999). *eCommerce Trust Study*. [Date accessed: 10/07/08].
<http://www.cheskin.com/cms/files/i/articles//17__report-eComm%20Trust1999.pdf>

CORRITORE, C.L.; KRACHER, B.; WIEDENBECK, S. (2003). "On-line trust: concepts, evolving themes, a model". *International Journal of Human-Computer Studies*. Vol. 58, pp. 737-758.

DHAMIJA, R.; TYGAR, J.D.; HEARST, M. (2006). "Why phishing works". *Conference on Human Factors in Computing Systems (CHI 2006)*. [Date accessed: 10/07/08].
<http://people.seas.harvard.edu/~rachna/papers/why_phishing_works.pdf>

DUTTON, W.H.; HELSPER, E.J. (2007). *The Internet in Britain 2007*. Oxford: Oxford Internet Institute.

FUKUYAMA, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press.

GOLBECK, J. (2006). "Trust on the World Wide Web: a survey". *Foundations and Trends in Web Science*. Vol. 1, iss. 2, pp.1-72.

GOLBECK, J.; PARSIA, B.; HENDLER, J. (2003). "Trust networks on the Semantic Web". *Proceedings of Cooperative Intelligent Agents*, Helsinki, 2003.
[Date accessed: 11/07/08].
<http://www.mindswap.org/papers/CIA03.pdf>

GOLBECK, J.; ROTHSTEIN, M. (2008). "Linking social networks on the Web with FOAF: a Semantic Web case study". *Proceedings of AAAI '08*. [Date accessed:11/07/08].
<http://www.cs.umd.edu/~golbeck/downloads/foaf.pdf>

HARDIN, R. (1996). "Trustworthiness". *Ethics*. Vol. 107, iss. 1, pp. 26-42.

HYUNH, T.D.; JENNINGS, N.R.; SHADBOLT, N.R. (2006). "An integrated trust and reputation model for open multi-agent systems". *Journal of Autonomous Agents and Multi-Agent Systems*. Vol. 13, pp. 119-154. [Date accessed:10/07/08].
<http://eprints.ecs.soton.ac.uk/12593/1/jaamas-dong.pdf>

LUHMANN, N. (1980). *Trust and Power*. Chichester: Wiley.

MOREAU, L.; GROTH, P.; MILES, S. [*et al.*] (2008). "The provenance of electronic data". *Communications of the ACM*. Vol. 51, iss. 4, pp. 52-58.

O'HARA, K. (2004). *Trust: From Socrates to Spin*. Duxford: Icon.

O'HARA, K.; ALANI, H.; KALFOGLOU, Y. [*et al.*] (2004). "Trust strategies for the Semantic Web". *Workshop on Trust, Security and Reputation on the Semantic Web*, Hiroshima, Japan. [Date accessed:11/07/08].
<http://eprints.ecs.soton.ac.uk/10029/>

O'HARA, K.; STEVENS, D. (2006). *inequality.com: Power, Poverty and the Digital Divide*. Oxford: Oneworld.

O'HARA, K.; SHADBOLT, N. (2008). *The Spy in the Coffee Machine: The End of Privacy As We Know It*. Oxford: Oneworld.

O'HARA, K.; TUFFIELD, M.; SHADBOLT, N. (2008). "Lifelogging: issues of privacy and identity with Memories for Life". *Workshop on Identity in the Information Society*, Arona, Italy, May 2008. [Date accessed:11/07/08].
<http://eprints.ecs.soton.ac.uk/15993/>

RICHARDSON, M.; AGRAWAL, R.; DOMINGOS, P. (2003). "Trust management for the Semantic Web". In: D. FENSEL, K. SYCARA, J. MYLOPOULOUS (eds.). *The Semantic Web – ISWC 2003*. Berlin: Springer, pp. 351-368.

SHADBOLT, N.; HALL, W.; BERNERS-LEE, T. (2006). "The Semantic Web revisited". *IEEE Intelligent Systems*. Vol. 21, iss. 3, pp. 96-101.

USLANER, E.M. (2002). *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.

WEITZNER, D.J.; HENDLER, J.; BERNERS-LEE, T. [*et al.*] (2005). "Creating a Policy-Aware Web: discretionary, rule-based access for the World Wide Web". In: E. FERRARI, B. THURAISINGHAM (eds.). *Web and Information Security*. Hershey, PA: Idea Group Inc. [Date accessed:11/07/08].
<http://www.w3.org/2004/09/Policy-Aware-Web-acl.pdf>

Recommended Reference:

O'HARA, Kieron; HALL, Wendy (2008). "Trust on the Web: Some Web Science Research Challenges".
In: Julià MINGUILLÓN (coord.). "Web Science" [dossier online]. *UOC Papers*. Iss. 7. UOC. [Date accessed: dd/mm/yy].
<http//www.uoc.edu/uocpapers/7/dt/eng/ohara_hall.pdf>
ISSN 1885-1541

Kieron O'Hara
Senior research fellow in Electronics and Computer Science at the University of Southampton (United Kingdom)
kmo@ecs.soton.ac.uk

Kieron O'Hara is a senior research fellow in Electronics and Computer Science at the University of Southampton and a fellow of the Web Science Research Initiative. He researches into the politics, philosophy and epistemology of technology. He is the author of several books, including: *Plato and the Internet* (2002); *Trust: From Socrates to Spin* (2004); *After Blair: Conservatism Beyond Thatcher* (2005); *The Referendum Roundabout* (2006); *inequality.com: Power, Poverty and the Digital Divide* (2006, with David Stevens); and *After Blair: David Cameron and the Conservative Tradition* (2007). He also wrote "A Framework for Web Science" (2006, with Tim Berners-Lee, Wendy Hall, James A. Hendler *et al.*), for the journal *Foundations and Trends in Web Science*, and edited *Advances in Knowledge Acquisition* (1996, with Nigel Shadbolt and Guus Schreiber). He also publishes frequently in popular journals and newspapers, and has appeared several times on radio and television.
Personal webpage: <http://www.ecs.soton.ac.uk/people/kmo>.

Wendy Hall
Professor of Computer Science
at the University of Southampton
(United Kingdom)
wh@ecs.soton.ac.uk

Wendy Hall is Professor of Computer Science at the University of Southampton, UK. She was Head of the School of Electronics and Computer Science (ECS) from 2002 to 2007. One of the first computer scientists to undertake serious research in multimedia and hypermedia, she has been at its forefront ever since. The influence of her work has been significant in many areas including digital libraries, the development of the Semantic Web and the emerging research discipline of Web Science. Her current research includes applications of the Semantic Web and exploring the interface between the life sciences and the physical sciences. She is a Founding Director, along with Professor Sir Tim Berners-Lee, Professor Nigel Shadbolt and Daniel J. Weitzner, of the Web Science Research Initiative.

In addition to playing a prominent role in the development of her subject, she also helps shape science and engineering policy and education. Through her leadership roles on national and international bodies, she has shattered many glass ceilings, readily deploying her position on numerous national and international bodies to promote the role of women in SET, and acting as an important role model for others.

She has recently been elected President of the Association for Computing Machinery (ACM), and is the first person from outside North America to hold this position. Until July 2008, she was Senior Vice President of the Royal Academy of Engineering, is currently a member of the UK Prime Minister's Council for Science and Technology, and is a founder member of the Scientific Council of the European Research Council. She was President of the British Computer Society (2003-4) and an EPSRC Senior Research Fellow from 1996 to 2002.

Personal webpage: <http://users.ecs.soton.ac.uk/wh/>

Universitat Oberta de Catalunya

uocpapers e-Journal on the Knowledge Society

http://uocpapers.uoc.edu

article

"Web Science" Dossier

E-learning from the Perspective of Web Science: Looking to the Future

Julià Minguillón

Submission date: September 2008
Accepted in: September 2008
Published in: October 2008

Abstract

This article shows the evolution that distance education has undergone recently from a new perspective, Web Science, which studies how the Web has evolved not only technologically but also socially and organisationally. A technological solution initially designed to share information, the Web is now present in all everyday activities and in every sphere – personal, academic and professional – and has changed the way we relate with one another, work and, obviously, access shared knowledge and learning. With the appearance of the Web, distance education has ceased to be a second option relegated to students not having the chance of accessing the university education system and is becoming common in the university education system, allowing students to take control of their lifelong learning process, both academic and professional, without the barriers of time or space. A number of factors have brought about this change – technological, methodological and organisational – but also social changes. Web Science studies how all of these changes are interrelated and their influence on such areas as the economy, leisure and education, the field of interest of this article. E-learning as the evolution of distance education is (or should be), therefore, a clear example of a Web Science case study, in which all of these aspects occur.

Keywords

e-learning, distance education, Web Science, virtual learning environments, Internet, Web

Resum

En aquest article es presenta l'evolució que ha sofert recentment l'educació a distància des d'una nova perspectiva, la ciència del web, que estudia com el web ha evolucionat tecnològicament, però també socialment i organitzativament. El web, una solució tecnològica inicialment pensada per a compartir informació, és avui dia present en totes les activitats quotidianes, en qualsevol àmbit –personal, acadèmic o professional–, i ha modificat la forma de relacionar-se, treballar i, evidentment, accedir al coneixement comú i aprendre. Amb l'aparició del web, l'educació a distància ha deixat de ser una segona opció relegada a estudiants sense la possibilitat d'accedir al sistema educatiu universitari i es va convertint en un fet comú en aquest sistema, on l'estudiant pren el control del procés d'aprenentatge al llarg de la vida, acadèmica i professional, sense barreres temporals o espacials. Diferents factors han afavorit aquest canvi: tecnològics, metodològics i organitzatius, però també socials. La ciència del web estudia com tots aquests canvis estan relacionats entre ells i la seva influència en àrees com l'economia, l'oci o l'educació, objecte d'interès d'aquest article. L'e-learning com a evolució de l'educació a distància és (o hauria de ser), doncs, un clar exemple de cas d'estudi de la ciència del web on es donen tots aquests aspectes.

Paraules clau

e-learning, *educació a distància*, Web Science, *entorns virtuals d'aprenentatge*, *internet*, *web*

Introduction

Distance education has undergone a great change since its beginnings, primarily because all of the elements involved have changed radically: the technology used, the educational and organisational models and, especially, the students themselves, who are more aware of the time and effort that studying (distance or face-to-face) represents and try to maximise their performance. Technology has enabled this change (Bates, 1995), yet as we will see below, there are many other factors that are influencing and determining the form that education at any level, but especially university education, will have in the not too distant future. As Bates (2008) stated, we are on the point of entering the fifth generation of distance education (Taylor,

1999), where the student will be the key element of every e-learning-based education process.

The concept of e-learning, used to indicate a learning process supported by the intensive use of technology, has changed with time, and nowadays there are in fact different definitions or ideas of what we mean by e-learning, depending on the context (education, business, etc.). The basic idea behind the concept of e-learning is that the teaching and learning process occurs through the action of a certain technology, specifically Information and Communication Technologies (ICT). E-learning is an extension of the previously-used concept of Computer-Based Training, which dates back to the 1980s, coinciding with the appearance of personal computers and the first multimedia systems, called *second-generation systems*. In fact, the word e-learning has been used both to describe self-learning multimedia courses distributed on a CD-ROM and for integrated tools in learning environments with primitive computers (Graziadei, 1993), which has led to its being a confusing concept used for any educational experience no matter how small the presence of ICT. Today, the most widely accepted meaning of e-learning coincides with the fourth generation described by Taylor (1999), where there is an asynchronous process that allows students and teachers to interact in an educational process expressly designed in accordance with these principles.

Despite this, to avoid confusion, today we prefer to speak of Internet-Based Learning or, better still, Web-Based Learning, for example, to explain the fact that distance education is carried out using the Internet, with the appearance of the virtual learning environment concept, a web space where the teaching and learning process is generated and supported (Sangrà, 2002). This entails overcoming the barriers of space and time of brick and mortar education (face-to-face) or of distance education using broadcasting and adopting a completely asynchronous model that allows access to education by many more users, at any level (including secondary education, but primarily university education and lifelong learning).

Under this paradigm, besides the purely virtual educational institutions, still few in number, increasingly more universities and secondary schools are adopting and incorporating virtual learning environments to complement the education of their students, albeit with diverse results (Curran, 2001). Adopting a purely virtual or blended model is not simply incorporating the necessary technology but also implementing methodological and organisational changes (Thomas *et al.*, 1998). Posting documents on a website does not automatically lead to a learning process, but the organisational and methodological challenges associated with technological change must be posed (Bates, 1995). In fact, some authors are critical of this process of adoption of virtual technology by the traditional universities, highlighting the failures of such initiatives as NYU Online and Cornell University, to name just two (Bang, 2006). Apart from this, there are also initiatives that have been successful and become benchmarks in university virtual teaching, such as the University of Phoenix and the UOC. The popularisation of proprietary and especially open tools for course and educational content management has been a key factor in the adoption of this technology (Boneu, 2007).

As regards non-regulated, or informal, learning, Internet users organise themselves around communities of interest in any subject, no matter how minor it may seem. Technology has simplified the creation of these communities, no matter how small and specialised they are. The fact of having search engines like Google allows any resource to be found; in fact, the problem is not finding resources on a subject but evaluating which resource is the most suitable for a specific use. This is just one more effect of the so-called *long tail*, described by Chris Anderson (2004) to refer to the business models based on Internet use for the dissemination of their products, such as Amazon, for example, which can have catalogues with millions of entries even though the majority of them only potentially interest one single user. In the same way, the Internet has allowed geographically disperse users to group themselves around common interests, no matter how minority they may be. Some of these communities of interest become genuine learning communities, where the users construct knowledge collectively, following informal learning models, sharing experiences and interacting using the tools provided by the community. Technology provides the necessary elements to motivate the students and make the most of their experience (Huitt, 2001). A good example of this fact is the DLESE[1] repository devoted to earth sciences, on the basis of a digital library with resources on the subject, focusing on a clear educational objective (Sumner *et al.*, 2001). It has been the Internet network that has enabled this radical change in access to information and the creation of communities around it, offering users a range of tools and services for communication, both synchronous and asynchronous.

The Internet Network

The Internet, the network of networks, is the medium that we all know as the World Wide Web, becoming popularised since the appearance of the first protocols (HTTP), which allowed the

1. Visit: <http://dlese.org>.

linking of documents located on other machines that were part of the network. These documents, written in HTML, permit the incorporation of text, images and links to other documents and are increasingly becoming true hypermedia documents, creating an enormous range of information locatable simply by the use of search engines such as Google. This is what is known as Web 1.0, a read-only web offering access to the available information. A recent estimate[2] points to more than 1.4 billion Internet users worldwide accessing over 170 million websites[3], which gives an idea of the huge size of it all. Having the suitable informational competences to find, filter and select the appropriate information is, therefore, a vital necessity for Internet users, and in this sense there are increasingly more tools that provide support to these needs. The Internet is evolving beyond this simple model of inter-linked hypertext documents towards a heterogeneous knowledge database in multiple formats, giving ever more importance to the concept of resource or content. Today, the Internet is a reflection of the real world in all its dimensions and is present in every aspect of the everyday life of its users, be it professionally, academically or personally.

Following its natural evolution, the Web is currently immersed in what is known as Web 2.0, a social movement that gives more power to the end users, as it is these who create, publish and manage their own content (Geser, 2007) through the use of very simple tools, such as blogs, wikis, flickr and YouTube, to name just a few. The Web 1.0 was reserved for experts and worked, more or less, in one direction, with a small number of people generating content for the great majority of users. Web 2.0 is a real social movement supported by the technology, which has democratised user participation on the net, going from being simple consumers of information to its producers, without having to be expert in any technology. The best example of this movement is Wikipedia,[4] a free and multilingual encyclopaedia constructed and maintained collaboratively by thousands of users worldwide and which, since its creation in 2001, has become one of the most visited websites in the world, with more than 10 million articles in 253 languages, and over two million in English.

The field of education does not escape this movement either. On the one hand, students are increasingly using Web 2.0 tools to work collaboratively as a common part of their learning process, which is more participative and student-centred, following the guidelines of the new European Higher Education Area. On the other, educational institutions and their users (teachers and students) are beginning to publish educational contents created by them in open form so that everyone can access them freely and at no cost. This movement, Open Educational Resources, has become a true show of strength by users, similar to the one represented by the free and open source software movement in its day (Geser, 2007).

In fact, both movements have implications in the educational sector, as for the first time it is possible to have a suitable technological platform to support a virtual learning environment and also the necessary contents that comprise the courses, ie, the educational offer. Despite this, there are other factors that are necessary to provide a more detailed analysis of the evolution of distance education using the Internet, following the analytical model proposed by what is known as Web Science.

The Web as Science

What we understand as the Web has also evolved over time since its invention (as we understand it today) in 1989, with Tim Berners-Lee being regarded as its creator. It began to become popular after 1993 with the appearance of the first search engines, such as Mosaic 1.0, while 1994 saw its mass adoption by users. Today, the Web is an everyday tool in any context (professional, academic and personal) and has become a "transparent" infrastructure in the sense that it is always present without the intervention of users, like electricity. The possibility of connecting using mobile devices through wireless networks has enhanced this sensation of immediacy.

The Web has evolved technologically, but above all it has had an impact on how users relate and how they use the Web for their everyday academic and professional activities. The concept of Web Science, used for the first time by Tim Berners-Lee *et al.* (2006), seeks to embrace all the technological, organisational and social aspects relating to the Web, from a multidisciplinary perspective, with the aim of understanding what the Web is and how it is used by its users, so to speak, and how this use establishes an interdependence between what the users do and the evolution of the Web. Web Science includes aspects ranging from the micro, such as the protocols and technology used to provide support to the Web, to the macro, such as the social phenomenon represented by blogging, for example, and its social implications. This multi-level approach allows any question relating to the Web to be focused using different methodologies, each one geared towards resolving a specific aspect of the phenomenon under study, as the Web is impossible to explain from an exclusively technological or social point of view. The Web is like a human being with its biology and ecology which

2. Visit: <http://www.internetworldstats.com/stats.htm>.
3. Visit:<http://news.netcraft.com/archives/web_server_survey.html>.
4. Visit: <http://www.wikipedia.org/>.

evolves constantly at a speed that hinders its understanding (Hendler *et al.*, 2008). From the Web emerges an intelligence, part artificial, part social, which is the driving force of these continuous changes (Salem, 2007).

A clear example of this is what is known as the Semantic Web. In 2004, Tim Berners-Lee said at the opening session of the 13th International Conference on the World Wide Web that the Web would be semantic or it would not be, in the sense that without a system that allows machines to extract information from the Web in a similar way to how humans do it, the net would be a huge information (or disinformation) dump from which it would be very difficult to get any use. It is necessary for all the resources (in the broad sense of the word) available on the net to be correctly described using the technology currently available, such as RDF. If the available information is well structured and labelled, it will be possible to retrieve it according to certain specific search criteria, permitting the location of the most suitable resources for a specific task or activity. In this respect, a well organised Web becomes a highly valuable educational space, since it enables two of the basic requirements of any learning process to be met: the existence of educational contents and the interaction with other users with the same interests. However, not only with regard to the users, but also with regard to automatic systems that locate and select resources according to the preferences and particularities of the users, essential tools for providing support to the users of virtual learning environments.

The future of the Web, called 3.0 by some authors (Spivack, 2006), involves achieving a higher degree of technological maturity, which will permit greater interoperability between all of its elements, including users, services and machines. Without intending to enter into philosophical disquisitions, Web 3.0 will incorporate a certain intelligence and semantics that will allow users to take a qualitative leap in its use (Salem, 2007). It is not a case of the Web becoming aware of itself (Tetlow, 2007, pp. 163-167), but of exponentially multiplying the possibilities that it offers users, facilitating their operations and anticipating their needs. In this sense, the concept of distance education using the Internet will also have to evolve, placing the student at the centre of the learning process and providing them with the adequate tools to achieve their short-, medium- and long-term aims. For example, reinforcing the feedback that the distance student receives or increasing the capacity to use educational resources.

E-learning As an Example of Web Science

It is clear, then, that the evolution that distance education has undergone since the appearance of the Internet has been radical, even though there is still a long way to go. Different movements and events have converged to change completely the traditional meaning of what was considered distance education, relegated to a second option for those who could not access the established education circuits. Often distance education has been seen only as a possibility for adults, without the time to attend master classes in the classic brick and mortar university model, which has led to a dangerous association between the distance education concept and the low quality of teaching, measured through the performance of the students, an association very often backed by the absolute lack of an educational and support model for the distance student, who is evidently highly prone to dropping out in a situation of complete isolation. The combination of limited technology with educational models based on the simple transmission of contents has been the reality to date of distance education that has only been of use to the most motivated students (Huitt, 2001).

However, fortunately, this perception is changing thanks to the intervention of highly diverse factors, including technological (bandwidth, mobile devices, wireless communications networks, free software, calculation capacity, etc.), educational (user-centred learning models), methodological (the new European Higher Education Area) and legal (open source). Yet above all, the perception held by students of the classic university model is also changing, which is striving not to appear outdated in the face of all of these new technologies. The fact that all the brick and mortar universities currently offer part of their training in blended or purely virtual modes is also evidence that the classic model is considered to be obsolete. Consequently, e-learning has radically changed the concept of distance education as its key elements have evolved:

1. Today there is a cheap and very powerful technology that enables instant and continuous connection to the Internet using mobile devices. There is also technology to be able to create very complex immersive educational simulations and scenarios that allow the learning process to be improved in complex areas such as medicine and engineering. This is what is called Technology-Enhanced Learning.
2. The Internet today is a virtual space where not only can we search for and obtain information of all kinds, but also create, modify and share it. Web 2.0 has given end users the power held until now by the contents creators, who held a monopoly on what could be found on the net. The do-it- yourself philosophy is increasingly becoming do-it-ourselves. This has a great impact on the student's learning process, getting them involved more.
3. The guidelines set by the new European Higher Education Area promote a competence-based student-centred

learning that is more homogeneous among universities, which allows the students to study and work in any country (in terms of the European sphere). This can lead to greater competition between universities as, through quality distance education, the students will not be tied to an option that limits them due to geographical reasons, for example. The distance universities now not only aim at a small number of adult students, but at their whole lifecycle, from when they enter the university to continuous professional development.

4. The business world and employment market increasingly demand well-trained professionals who can adopt all technological and methodological changes quickly. Distance education is a key tool in ensuring a lifecycle of the student consistent with their professional development. In fact, businesses have been promoters of e-learning, now is the time to share experiences and needs with the university.
5. Students are aware of the changes fostered by the above factors and they demand greater participation in the learning process, taking decisions in this respect, at all times within the established framework of course, in line with their personal objectives and with a clearly professional interest. The net is a space where the different identities of each individual (academic, professional and personal) can converge, a clear example of this being communities such as MySpace and Facebook. In fact, we now speak of a generation of digital natives who do not need to learn to use a technology but where the use of technology is a competence they have already acquired. This group is not afraid of technology but demands it, rejecting experiences based on obsolete models.
6. Finally, there begins to be a significant critical mass of open educational contents that are available for use on open e-learning platforms under licences that allow the users to create and share the contents created by them freely (Megías, 2007). Although this point is still at a very early phase (more with regard to open contents than the platforms, which are more advanced), it is a subject that has received a great deal of attention in recent years (Minguillón, 2007). This possibility enables educational institutions to offer training via e-learning and disseminate their knowledge.

All of these aspects mean that distance education over the Internet is making a qualitative and quantitative leap. The following table shows the equivalences between the different concepts that have evolved (and continue to evolve) with and via the Web, all of which are related to distance education and resource management over the Internet.

Table 1. Equivalences of concepts related to distance education and resource management over the Internet

past/present	present/future
One-way dissemination (broadcasting)	Two/multi-way dissemination
Push model (the user receives)	Pull model (the user chooses)
Producer-consumer model	Create-mix-share model
HTML pages	CMS, content managers
Contents	Activities
Knowledge acquisition	Competence development
Portfolio	e-portfolio
Classroom, blackboard	Virtual classroom
Photocopies, notes	PDF, e-books
Exercises, examples	Educational resource repositories
Final exams	Continuous assessment
Modem	ADSL, Wi-Fi
e-mail	Instant messaging
FTP	P2P
Forums, noticeboards	Blogs, wikis
Copyright	Copyleft, Creative Commons
Institutional contents	Collective contents
Institutional web portals	Individual web spaces
Communities of interest	Learning communities
Virtual learning environments	Personal learning environments

The Future of E-learning

As Taylor (1999) describes, the fifth generation of e-learning systems will have to make the most of the possibilities that the Internet offers (or will offer), placing the emphasis on the student, centre of the learning process. Such a visionary as Stephen Downes (2005) already points to the use of tools that favour social learning, creating and sharing knowledge collaboratively in an increasingly informal scenario, yet one that is increasingly controlled and supported by the technology.

In an ideal future scenario, the student uses one or more devices to access the Internet, where all their academic, professional, etc. records are stored, as are all their preferences and particularities. The student, who wants to attain educational objectives, needs to acquire and develop a series of competences through an itinerary designed especially for them. The virtual learning environment guides the student on this itinerary, providing them with all the resources that they need at all

times, either automatically or on demand. The system records all the student's actions in order to obtain valuable information for the institution, improving the design of itineraries, usability aspects of the user interface, detecting problematic educational resources, measuring the student's degree of interaction with the system and the other students and lecturers, etc. This record also serves to assess the student, as the learning process is considered to be as or more important than the results obtained, storing information in their profile on the activities carried out (information searches, interaction, etc.) and the results obtained, with the aim of improving the degree of personalisation of their learning process. When students have a problem or query, the system helps them to resolve it or puts them in touch with other people who can help them, providing continuous feedback that prevents the feeling of isolation felt by online students.

We are still a long way from the above. However, the world's top distance universities today are leading the way in research in fields related to e-learning, such as the Open University (United Kingdom), the Open Universiteit (The Netherlands), and the UOC with its Virtual Campus. Such subjects as personalisation of the learning process, semantic repositories of learning objects and intelligent tutors offer solutions to each of the problems to be resolved in this ideal scenario. In turn, all of this technology generates new needs, such as open source for educational contents, authorship tools, business models, quality measures, privacy and security aspects, etc., which clearly shows that e-learning needs a formal multidisciplinary and multi-level approach like the one proposed by Web Science.

Conclusions

Virtual learning environments are a reality used by educational and business institutions of all kinds, from small schools to large universities, or training departments in companies. These virtual environments partially reproduce the elements present in the teaching and learning process through a series of processes, services and resources. The evolution of content management systems and the learning process towards genuine virtual learning environments is coming about, in part, due to the pressure that initiatives such as the new European Higher Education Area are exercising on educational institutions, but also so as to meet the new educational needs of their users, who demand more active, collaborative and personalised learning, where the student has control over most of the learning process. This is what some authors have called *fifth-generation e-learning*.

In the not too distant future, users will access their work space using mobile devices, where all their identities (academic, professional and personal) will converge, with the borders between them becoming ever more blurred, especially for users immersed in lifelong learning strongly related to their professional development but also to their private interests. As the technology evolves and the mobile devices become increasingly more powerful and allow for huge quantities of information to be accessed and stored instantly, the concept of virtual learning environment will cease to be a model based on a centralised system that provides all the necessary processes, services and resources and become a virtual space that every student will carry with them at all times. The centralised model will continue to exist but only for enormous resources such as institutional repositories, digital libraries and super-computing centres, which the student will access from their device.

Today we are at the point where the necessary conditions have arisen to take a step forward in the concept of e-learning, as described by Web Science. A combination of key factors at micro and macro levels have come together to make what was known as distance education evolve. E-learning has replaced this concept through a technology, the Web, and has also changed the perception that users have of it, obliging every educational institution to adopt it in order not to be left behind in the race towards quality lifelong education.

Unfortunately, this scenario faces various problems. One possible one is the digital divide posed by this new scenario, based on a tool such as the Internet, which has different points of access at very diverse speeds and costs. This barrier also exists in the form of the knowledge needed, as not everyone has received the adequate training in working with ICT. There is also still some reticence in terms of the quality of a form of teaching that continues to be associated with a secondary option. Despite this, e-learning is unstoppable because the Web is unstoppable, with a technology that is evolving rapidly and users who adopt and adapt it to their needs in a continuous circle of improvement.

Reference

ANDERSON, C. (2004). "The Long Tail" [online]. *Wired,* Vol. 12, no. 10. [Date accessed 07/07/2008]. <http://www.wired.com/wired/archive/12.10/tail.html>

BANG, J. (2006). "eLearning Reconsidered. Have e-learning and Virtual Universities Met the Expectations?" [online]. *E-learning europa*. [Date accessed 27/08/2008]. <http://www.elearningeuropa.info/directory/index.php?page=doc&doc_id=7778&doclng=6>

BATES, T. (1995). *Technology, Open Learning and Distance Education*. London: Routledge.

BATES, T. (2008). "Transforming Distance Education through New Technologies". In: T. EVANS, M. HAUGHEY and D. MURPHY (eds.). *The International Handbook of Distance Education.* Bingley, UK: Emerald Press.

BERNERS-LEE, T. [*et al.*] (2006). "Creating a Science of the Web". *Science.* Vol. 313, no. 5788, pp. 769-771.

BONEU, J.M. (2007). "Plataformas abiertas de e-learning para el soporte de contenidos educativos abiertos". In: J. MINGUILLÓN (coord.). "Contenidos educativos en abierto" [online monograph]. *Revisa de Universidad y Sociedad del Conocimiento (RUSC).* Vol. 4, no 1. UOC. [Date accesssed: 18/08/2008]. <http://www.uoc.edu/rusc/4/1/dt/esp/boneu.pdf>

CURRAN, C. (2001). "The Phenomenon of On-Line Learning". *European Journal of Education.* Vol. 36, no. 2, pp. 113-132.

DOWNES, S. (2005). "E-learning 2.0" [online]. *Elearn.* Vol. 10. [date accessed: 26/08/2008]. <http://doi.acm.org/10.1145/1104966.1104968>

GESER, G. (ed.) (2007). *Open Educational Practices and Resources. OLCOS Roadmap 2012* [online]. ISBN 3-902448-08-3. [date accessed: 08/07/2007]. <http://www.olcos.org/cms/upload/docs/olcos_roadmap.pdf>

GRAZIADEI, W.D. (1993). "Virtual Instructional Classroom Environment in Science (VICES)" [online]. *Research, Education, Service & Teaching (REST).* [Date accessed: 18/08/2008]. <http://www.cni.org/projects/netteach/1993/prop01.html>

HENDLER, J. [*et al.*] (2008). "Web Science: An Interdisciplinary Approach to Understanding the Web". *Communications of the ACM.* Vol. 51, no. 7, pp. 60-69.

HUITT, W. (2001). "Motivation to Learn: An Overview" [online]. *Educational Psychology Interactive.* Valdosta, GA: Valdosta State University. [Date accessed: 18/07/2008]. <http://chiron.valdosta.edu/whuitt/col/motivation/motivate.html>

MEGÍAS, D. (2007). "SELF: A Collaborative Platform to Share and Create Free Contents about Free Software and Open Standards" [online]. In: *Proceedings of the PKP Scholarly Publishing Conference*, Vancouver, Canada. [Date accessed: consultation 27/08/2008]. <http://scholarlypublishing.blogspot.com/2007/07/self-collaborative-platform-to-share.html>

MINGUILLÓN, J. (2007). "Contenidos educativos en abierto" [online monograph]. *Revista de Universidad y Sociedad del Conocimiento (RUSC).* Vol. 4, No. 1. UOC. [date accessed: 18/08/2008]. <http://www.uoc.edu/rusc/4/1/dt/esp/monografico.pdf>

SALEM, A.-B. M. (2007). "Web Intelligence: A New Paradigm for Virtual Communities and Web Science". In: *Proceedings of the 8th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services*, pp. 33-35.

SANGRÀ, A. (2002). "A New Learning Model for the Information and Knowledge Society: The Case of the UOC" [online]. *International Review of Research in Open and Distance Learning.* Vol. 2, no. 2. [Date accessed: 18/08/2008]. <http://www.irrodl.org/index.php/irrodl/article/view/55>

SPIVACK, N. (2006). "The Third Generation Web Is Coming" [online]. *KurzweilAI.* [date accessed: 25/08/2008]. <http://www.kurzweilai.net/meme/frame.html?main=/articles/art0689.html>

SUMNER, T.; DAWE, M. (2001). "Looking at Digital Library Usability from a Reuse Perspective". In: *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, New York, NY, pp. 416-425.

TAYLOR, J.C. (1999). "Distance Education: the Fifth Generation". In: *Proceedings of the 19th ICDE World Conference on Open Learning and Distance Education*, Vienna, Austria.

TETLOW, P. (2007). *The Web's Awake. An introduction to the Field of Web Science and the Concept of Web Life.* New Jersey: Wiley and Sons.

THOMAS, P.; CARSWELL, L.; PRICE, B. [*et al.*] (1998). "A Holistic Approach to Supporting Distance Learning Using the Internet: Transformation, not Translation". *British Journal of Educational Technology.* Vol. 29, no. 2, pp. 149-161.

Recommended Reference:

MINGUILLÓN, Julià (2008). "E-learning from the Perspective of Web Science: Looking to the Future". In: "Web Science" [dossier online]. *UOC Papers.* Iss. 7. UOC. [Date accessed: dd/mm/yy]. <http//www.uoc.edu/uocpapers/7/dt/eng/minguillon.pdf>
ISSN 1885-1541

Julià Minguillón
Assistant Director of the IN3 (UOC)
jminguillona@uoc.edu

Julià Minguillón gained his PhD from the Autonomous University of Barcelona in September 2002, where he is associate lecturer in the Information and Communications Engineering Department. In 2001, he joined the UOC as a lecturer in the IT, Multimedia and Telecommunications Department, where he carries out teaching work in the areas of programming, languages and compilers, graphic computing, statistics and data-mining. He has taken part in the creation of teaching resources on object-oriented programming, information structure, abstract data types and compilers. He leads the UOC NET2LEARN research group, where his interests include the description and standardisation of educational contents and the learning process using ontologies, semantic repositories of learning objects for personalisation using training itineraries, and modelling the behaviour of users in a virtual learning environments using Web-mining techniques. He has taken part in a number of national and international projects related to e-learning, among which he has co-managed the OLCOS project on open educational contents. He currently manages the PERSONAL(ONTO) project on personalisation and the E-MATHPP project on learning object repositories. He organised the second SPDECE 2005 symposium on reusable learning objects and has played an active role in the UOC UNESCO Chair in E-Learning since 2006.