

<http://uocpapers.uoc.edu>

article

Dossier «*Web Science*, la ciència del web»

La ciència, el web i la ciència del web

Daniel Riera

Data de presentació: setembre de 2008

Data d'acceptació: setembre de 2008

Data de publicació: octubre de 2008

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

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

La inèrcia del web

Segons la primera llei de Sir Isaac Newton (1643-1727), «tot cos lliure, sobre el qual no actua cap força, manté el seu estat de moviment, sigui en repòs o sigui en moviment rectilini uniforme». Aquesta primera llei rep el nom de llei d'inèrcia¹ i prova d'explicar l'actitud dels sistemes davant del moviment: si estan aturats, s'esforçaran per romandre aturats, i si es mouen, tractaran de continuar el moviment. Així, depenent de l'objec-

te (o sistema), podem imaginar-nos, sense aprofundir més en la física que hi ha darrere, que aquest objecte presentarà més o menys «resistència» al canvi. D'aquesta manera, si tenim una pilota de futbol i una bala de canó de la mateixa mida, totes dues parades l'una al costat de l'altra, sembla que serà més senzill imprimir moviment a la primera que no pas a la segona. Igualment, serà més fàcil aturar un camió que circula sense control a 1 kmh⁻¹ que una motocicleta, molt més petita, que s'apropa a 200 kmh⁻¹.

1. *Inertia*, en llatí, vol dir «oci, mandra».

Quan Sir Timothy John Berners-Lee va proposar l'any 1989 connectar la idea d'hipertext amb les de TCP i de DNS, estava posant a l'abast de tothom (ja que la infraestructura existia prèviament) el que avui dia anomenem «el web» (o el *World Wide Web*). Amb la idea de trobar una manera de facilitar la compartició i l'actualització immediata d'informació entre científics, es creava una xarxa que permetia veure continguts llunyans com si fossin a l'ordinador propi. A partir d'aquest moment, i amb aquesta petita empenta, una minsa bola de neu va començar a rodolar.

En no gaire temps (menys de dues dècades), el web ha anat agafant inèrcia. La incorporació, primer, de l'àmbit acadèmic, i, posteriorment, de l'empresarial i el particular, ha anat augmentant el volum de l'invent. A més, l'increment de les prestacions i la baixada dels preus, relacionats amb la ràpida evolució del maquinari, així com la incorporació de noves idees molt atractives al web, l'han fet madurar i enriquir-se molt ràpidament.

Si en els primers moments es veia la xarxa com la connexió d'un conjunt de continguts estàtics (o quasi estàtics), cal dir que aquests continguts s'han anat dinamitzant cada cop més. La xarxa ha passat a dir-se «web 2.0»² i ara sembla que es vol convertir en 3.0 (Markoff, 2006). La versió actual, la segona, aposta per la creativitat, la compartició de la informació (no només entre els científics) i la col·laboració mitjançant xarxes socials, wikis o blocs, entre d'altres. Molts d'aquests components ja existien des dels inicis del web, tot i que no eren coneguts ni usats tan extensament. La versió futura (tercera) proposa, principalment, la transformació en l'ús de la informació de la xarxa cap a un web semàntic (Berners-Lee *et al.*, 2001), amb totes les dades interconnectades i accessibles globalment, i amb motors de cerca amb un fort component d'intel·ligència artificial. A més, els continguts principalment seran oberts, així com les identitats dels usuaris, que han de ser portables a qualsevol màquina connectada a la xarxa.

L'entitat que ha assolit el web, a partir d'aquella empenta inicial que va rebre per part de Tim Berners-Lee aviat farà 20 anys, és immensa. La gran expansió va començar el 1995, quan ja es van comptar uns 16 milions d'usuaris. El 2001 ja eren 400 milions, i el 2005 es va arribar al primer bilió d'usuaris. S'espera que cap al 2010 se superin els dos bilions.³

Aquests dos creixements (qualitatiu i quantitatiu) fan que la inèrcia acumulada sembli incontrolable i que augmenti de manera exponencial. Una bona prova n'és que, des del punt de

vista docent, el que fa uns quants anys es podia resumir en un parell de lliçons en una assignatura d'un grau d'Informàtica, actualment ha assolit un volum que podria formar un grau complet tot per si sol.

Per tot això, el 2 de novembre de 2006, quatre investigadors del MIT i de la Universitat de Southampton van fundar la Web Science Research Initiative (WSRI).⁴ En la presentació⁵ que van fer van argumentar la necessitat d'un nou camp interdisciplinari –que uneixi matemàtiques, informàtica, psicologia, economia, dret, etc.– per a analitzar el que és i el que passa al web i al seu voltant. Un camp que mostri les vulnerabilitats, les interaccions que cal millorar, les implicacions que les decisions a escala microscòpica tenen en el disseny global, etc. Una ciència que treballi conjuntament els aspectes tècnics i els socials. A més, es proposa que al capdavant d'aquesta nova àrea hi hagi gent especialitzada, que sigui competent en aquest ampli ventall d'aspectes tan diferents *a priori*. Això requereix l'aparició d'oferta docent dedicada específicament a la *Web Science*.

La Web Science Research Initiative

El nombre de trobades internacionals relacionades d'una manera o una altra amb el web és molt gran. Tot i així, no és fins a la formació de la WSRI que s'ha començat a tractar el web de manera global i multidisciplinària. Des del seu naixement, la WSRI ha intentat fer conèixer la seva proposta mitjançant publicacions i presentacions en congressos i trobades internacionals. Ha provat de reunir experts dels diferents àmbits per començar a compartir tots la idea de «ciència del web».

Per tal d'aconseguir-ho, el mateix any 2005, als quarters generals de la British Computer Society⁶ a Londres, es va organitzar el Web Science Workshop: The Emerging Science of the Web,⁷ una invitació a la qual van acudir 21 participants. El tema central de la trobada va ser l'exploració dels desafiaments crítics del futur del *World Wide Web* (Hendler, 2005). Aquesta va ser la primera discussió formal global sobre la *Web Science*.

Posteriorment, s'han organitzat *workshops* dedicats exclusivament a aquest tema dins de conferències reconegudes, com ara la dissetena⁸ i divuitena⁹ conferències internacionals sobre el *World Wide Web*. En aquests tallers es demanava als participants que s'avancesin al futur del web i que discutissin sobre

-
2. No a tothom li agrada aquest nom. Per exemple, el mateix Tim Berners-Lee no el troba apropiat. De totes maneres, sembla àmpliament acceptat, en general, per la comunitat.
 3. Vegeu: <<http://www.internetworldstats.com/stats.htm>>.
 4. Vegeu: <<http://webscience.org/>>.
 5. Vegeu: <<http://web.mit.edu/newsoffice/2006/wsri.html>>.
 6. Vegeu: <<http://www.bcs.org/>>.
 7. Vegeu: <<http://www.cs.umd.edu/~hendler/2005/WebScienceWorkshop.html>>.
 8. Vegeu: <<http://www2008.org/>>.
 9. Vegeu: <<http://www2009.org/>>.

com aquest web es pot observar, analitzar i influenciar. Igualment, la conferència Hypertext 2008 (HT2008)¹⁰ ha inclòs el *workshop* Web Science: Collaboration and Collective Intelligence,¹¹ on s'ha parlat de noves formes de col·laboració a la xarxa, de la seva estructura i del comportament dels usuaris, principalment.

Entre la realització d'aquests *workshops*, s'han portat a terme moltes altres xerrades i presentacions que han servit a la WSRI i els seus components per a començar a propagar la seva proposta (Weitzner *et al.*, 2007; Hall, 2008; Hendler *et al.*, 2008). Igualment, alguns dels pares i mares de la WSRI han estat entrevistats o han escrit per a diaris d'interès general (Cellan-Jones, 2008; Lohr, 2008) o per a publicacions de l'àmbit científic (Schneiderman, 2007; Hendler *et al.*, 2008b). Al Web Science Curriculum Workshop, del setembre de 2008, es van debatre principalment els objectius de la iniciativa que estan relacionats amb la docència de la *Web Science*, i per a un futur proper ja s'ha planificat la First International Web Science Conference, que tindrà lloc a Atenes el març de 2009.

Una ràpida consulta a Google ens retorna, per a la cadena «*Web Science*», uns tres milions i mig de pàgines que ja fan referència a aquest tema. Entre aquestes pàgines, a part dels congressos, els articles i les entrevistes ja esmentats, trobem entrades a blocs en diversos idiomes (Saravanan, 2007; Peña, 2007; Ragget, 2008), notícies i articles en revistes i diaris en línia (Biever, 2006; LaMonica, 2006), així com alguna definició a llocs wiki. La majoria d'aquests articles i entrades a blocs són de novembre de 2006 i es limiten a fer conèixer l'anunci inicial de la WSRI.

Un dels darrers (i alhora dels primers) texts publicats és el provocatiu llibre *The Web's awake* (Tetlow, 2008), en què es presenta el web com una nova forma de vida que ja no podem controlar. D'aquesta manera, es descriuen noves característiques i comportaments del web que no hem creat nosaltres, sinó que han anat sorgint per si mateixos.

Com a curiositat, cal dir que la definició de *Web Science* a la Viquipèdia¹² és una mica pobra. Seguint l'esperit del web 2.0, són els usuaris els que haurien de completar aquest contingut.

A Framework for Web Science

Atesa la joventut de l'àrea, s'ha escrit poc sobre *Web Science*. De les poques referències que trobem, *A framework for Web Science* (Berners-Lee *et al.*, 2006) és possiblement la més completa. És el llibre¹³ amb què la WSRI ha provat de fixar els fonaments de la idea de *Web Science*. Tracta d'alguns dels temes centrals en el

desenvolupament de la xarxa, sense aprofundir en detalls excessivament tècnics. Atesa la quantitat i la varietat de coneixements que hi ha entorn del web, no sembla fàcil recollir-los en un sol monogràfic i de manera homogènia. Per això el llibre no prova de ser un recull tècnic d'aquests coneixements, sinó una introducció més aviat divulgativa dels temes principals que formarien la *Web Science*.

Així, després d'una breu introducció, es presenten la ciència i l'enginyeria que hi ha darrere, l'anàlisi de la xarxa, els aspectes socials que implica i, finalment, els temes legals, de governança, de seguretat i d'estàndards. Es fa un esforç important per establir la connexió entre temes força diversos i que fins ara s'havien analitzat de manera aïllada o, si més no, bastant independent. Tot i que aquests temes no queden totalment entrelaçats, sí que es comença a veure una estructura de la nova ciència que es proposa.

Alguns aspectes sobre els quals es fa èmfasi en el llibre són els relacionats amb l'organització de la xarxa, tant des del punt de vista social com de continguts i infraestructures o recursos. Així, se subratlla la importància d'un bon disseny de l'arquitectura de la xarxa. Igualment, s'aposta per l'organització i l'accés intel·ligent als continguts (o informació) mitjançant l'ús d'etiquetes, del web semàntic, etc. Es descriu la importància d'utilitzar models matemàtics per a l'anàlisi del web, que n'ha de permetre la millora contínua. D'altra banda, s'observen les implicacions dels contextos social, cognitiu i moral del web sobre les seves decisions d'enginyeria. Finalment, es pregunta sobre quin tipus de regulació s'ha d'aplicar al web i als seus usuaris. Això inclou principalment els temes relacionats amb la seguretat de l'usuari (p. ex., privadesa i identitat), els mals comportaments, la importància de l'acceptació per part de tothom dels estàndards i la cerca de l'estructura de govern de la xarxa.

La mancança principal que presenta el text és que l'heterogeneïtat dels temes de què tracta fa difícil un aprofundiment homogeni en la disciplina en si mateixa. A vegades el text sembla que es limita a ser una revisió d'una col·lecció d'articles ben escollits. Possiblement hauria estat més interessant si s'hagués fet un esforç més gran de coordinació dels diversos textos. Tot i així, és una obra valuosa que estableix els fonaments de la *Web Science*. Tracta de temes diversos, però força representatius del que actualment és el web i de les direccions que pren. Sembla estrany, això sí, que tot i tractar de tants de temes no s'hi pugui trobar ni la més mínima menció de la influència del web en l'ensenyament ni de conceptes com *e-learning* o educació a distància, quan és una de les àrees on el desenvolupament del web ha influït més.

10. Vegeu: <<http://www.sigweb.org/ht08/>>.

11. Vegeu: <<http://users.ecs.soton.ac.uk/dem/workshops/webscience08/>>.

12. Vegeu: <http://en.wikipedia.org/wiki/Web_science>.

13. De fet, no és exactament un llibre en si mateix, sinó el primer número de la revista dedicada a la ciència del web: *Foundations and Trends in Web Science*.

Contribucions acadèmiques

Un dels objectius de la WSRI, si no a curt termini, a mitjà termini, és el disseny conjunt entre el MIT i la Universitat de Southampton d'un grau de *Web Science*. Aquest grau avui dia encara no s'ofereix i no hi ha anuncis sobre quan s'ha de posar en marxa o dels avenços que s'han fet fins ara.

Una iniciativa propera a la de la ciència de la web és la de les *iSchools* (o *Schools of Information*),¹⁴ presentada el setembre de 2005 a la Universitat Estatal de Pennsilvània. Aquesta proposta difereix de la *Web Science* en el fet que se centra més en les relacions entre la informació, la tecnologia i les persones, independentment de la plataforma (que no ha de ser necessàriament el web). També té una oferta formativa relacionada amb el web l'Oxford Internet Institute,¹⁵ tot i que principalment va adreçada a màsters i doctorats. Igualment, a partir del concepte de *thread* que proposa la Georgia Tech,¹⁶ s'estableix la possibilitat que un estudiant es defineixi un currículum que, si bé encara no hi és gaire proper, es comenci a acostar al que seria un grau de *Web Science*. De totes maneres, aquest *thread* seria més proper al que es coneix actualment com a especialització dins d'un grau, que no pas a un grau en si mateix.

Finalment, en la proposta del futur grau d'Informàtica de la Universitat Oberta de Catalunya s'ha inclòs una assignatura optativa dedicada a la *Web Science*, mentre es prepara un postgrau sobre el tema.

Conclusions

A partir de la proposta de la WSRI i de la informació més fàcilment assolible, actualment, sembla que la *Web Science* es troba encara en una etapa inicial. Els seus membres segueixen fent esforços per donar a conèixer la proposta i fa la impressió que el tema central de discussió encara queda una mica difús. No sembla fàcil, és clar, definir amb precisió què entra dins d'aquest nou àmbit, què en queda fora, i quines són les interaccions entre els subàmbits.

Tot això no facilita el disseny de les noves titulacions associades. De moment, el MIT i la Universitat de Southampton suposadament preparen el primer grau de *Web Science*, mentre la resta d'universitats resten a l'espera. El currículum d'aquest nou grau ha de quedar ben consensuat entre un gran nombre de departaments d'una universitat, cosa que en complica el disseny.

Tot i això, la proposta només té dos anys de vida i sembla que s'estan fent força esforços per tirar-la endavant i definir bé el nou àmbit. La importància que ha d'assolir el web en la societat del se-

gle XXI es pot equiparar al que van significar la penicil·lina, la impremta, la màquina de vapor o l'electricitat en segles anteriors. La ciència del web és primordial per a assegurar-nos la formació de professionals capaços de conèixer i influenciar la xarxa i el que l'envolta.

Referències bibliogràfiques

- BERNERS-LEE, T.; HAENDLER, J.; LASSILA, O. (2001). «The Semantic Web» [article en línia]. *Scientific American*. [Data de consulta: 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, núm. 1.
- BIEVER, C (2006). «Web pioneers call for new "web science" discipline» [article en línia]. *New Scientist Tech*. [Data de consulta: 20/08/2008].
<<http://technology.newscientist.com/article/dn10438-web-pioneers-call-for-new-web-science-discipline.html>>
- CELLAN-JONES, R. (2008). Article i entrevista a la Dra. Wendy Hall [en línia]. BBC [Data de consulta: 20/08/2008]
<http://www.bbc.co.uk/blogs/technology/2008/03/what_is_web_science.html>
- HALL, W. (moderadora) (2008). *Does the Semantic Web Need Web Science?* [en línia] Panell debat a la «5th European Semantic Web Conference». Tenerife. [Data de consulta: 20/08/2008].
<http://videolectures.net/eswc08_hall_dsw/>
- HENDLER, J. (2005). Document resum del *Web Science Workshop. The Emerging Science of the Web* [en línia]. [Data de consulta: 20/08/2008].
<<http://www.cs.umd.edu/~hendler/2005/WebSciReport.pdf>>
- HENDLER, J.; BERNERS-LEE, T.; HALL, W. [et al.] (2008). Presentació del *Tetherless World Constellation: The Future of the World Wide Web* [en línia]. [Data de consulta: 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 (pàg. 60-69).
- LaMONICA, M. (2006). «Berners-Lee, universities launch "Web Science" initiative» [article en línia]. *c/net News.com*. [Data de consulta: 20/08/2008]
<http://news.cnet.com/Berners-Lee,-universities-launch-Web-science-initiative/2100-1025_3-6132016.html?hhTest=1, 2 de novembre 2006>

14. Vegeu: <<http://www.ischools.org/oc/>>.

15. Vegeu: <<http://www.oii.ox.ac.uk/>>.

16. Vegeu: <<http://www.gatech.edu/>>.

<http://uocpapers.uoc.edu>

La ciència, el web i la ciència del web

- LOHR, S. (2006). «Group of University Researchers to Make Web Science a Field of Study» [article en línia]. *The New York Times*. [Data de consulta: 20/08/2008].
<<http://www.nytimes.com/2006/11/02/technology/02compute.html>>
- MARKOFF, J. (2006). «Entrepreneurs See a Web Guided by Common Sense» [article en línia]. *New York Times*. [Data de consulta: 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* [article en línia (bloc)]. [Data de consulta: 20/08/2008].
<<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* [article en línia (bloc)]. [Data de consulta: 20/08/2008].
<<http://people.w3.org/~dsr/blog/?p=7>>
- SARAVANAN (2007). *Web Science: Web is a network of people* [article en línia (bloc)]. [Data de consulta: 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 (pàg. 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). Resum del *workshop Toward More Transparent Government Workshop on eGovernment and the Web* [en línia]. [Data de consulta: 20/08/2008].
<<http://www.w3.org/2007/06/eGov-dc/summary>>

Citació recomanada:

RIERA, Daniel (2008). «La ciència, el web i la ciència del web». A: Julià MINGUILLÓN (coord.). «*Web Science: la ciència del web*» [dossier en línia]. *UOC Papers*. Núm. 7. UOC. [Data de consulta: dd/mm/aa].
<<http://www.uoc.edu/uocpapers/7/dt/cat/riera.pdf>>
ISSN 1885-1541



Aquesta obra està subjecta a la llicència Reconeixement-NoComercial-SenseObraDerivada 2.5 Espanya de Creative Commons. Podeu copiar-la, distribuir-la i comunicar-la públicament sempre que n'especifiqueu l'autor i la revista que la publica (*UOC Papers*); no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/2.5/es/deed.ca>.

**Daniel Riera**

Director del Programa d'Enginyeria Informàtica dels Estudis d'Informàtica, Multimèdia i Telecomunicació (UOC)
drierat@uoc.edu

Doctor enginyer en Informàtica per la Universitat Autònoma de Barcelona (UAB) l'any 2006, màgister en Tècniques Avançades d'Automatització de Processos el 2001 per la UAB i enginyer d'Informàtica, també per la UAB, l'any 1998.

Becat pel Ministeri amb una beca de Formació de Personal Investigador (FPI), passa un any a l'Imperial College of London, com a visitador acadèmic, treballant a l'ICParc, al departament de planificació i control de recursos. Allà rep formació en programació amb restriccions i optimització, i col·labora en projectes pel RAC d'Anglaterra i Railtrack.

Ha fet docència a la UAB entre el setembre de 1998 i l'agost de 2005, i a partir de setembre de 2007. El setembre de 2005 entra a formar part de la plantilla de professors dels Estudis d'Informàtica, Multimèdia i Telecomunicació (EIMT) de la UOC. Des de setembre de 2006, és director acadèmic del postgrau de Bioinformàtica que actualment ofereix la UOC. Les seves activitats de docència es concentren principalment en la programació i la bioinformàtica. A partir del gener de 2007 s'incorpora com a director al Programa d'Enginyeria Informàtica dels EIMT.

Els seus àmbits de recerca principals inclouen el modelatge de sistemes discrets fent servir xarxes de Petri, l'optimització amb tècniques de programació amb restriccions. És autor de diversos articles en revistes i congressos, tant estatals com internacionals. Ha participat en projectes i convenis de recerca com a membre de l'equip de recerca. Ha estat investigador del LOGISIM, Centre de Simulació i Optimització de Sistemes Logístics, de la Xarxa de Centres de Suport a la Innovació Tecnològica (XIT) del CIDEM. Actualment pertany al grup de recerca sobre enginyeria del programari de la UOC (GRES-UOC), on investiga la verificació mitjançant l'ús de programació amb restriccions de models fets amb UML+OCL.