

The Emergence of Web Science



BEBO WHITE

SLAC

NATIONAL ACCELERATOR LABORATORY

CERN

OCTOBER 2008

Caveats

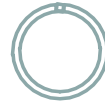


- I am not a member or researcher with the Web Science Research Institute (WSRI)
- Opinions about Web Science are my own

Web Science Research Institute (WSRI)



- Announced in November 2006
- A collaboration between MIT and the University of Southampton
- Stated purpose is “to bridge and formalize the social and technical aspects of collaborative applications running on large-scale networks like the Web.”
- “Brings together academics, scientists, sociologists, entrepreneurs and decision makers from around the world. These people will create the first multidisciplinary research body to examine the Web and offer the practical solutions needed to help guide its future use and design.”
- WebSci’09 – Athens, 3/18-20/2009



Before We Ask “What is Web Science?”

“What is ‘The Web?’” (1/2)



Wikipedia is sustained by people like you. Please [donate](#) today.

[article](#) [discussion](#) [edit this page](#) [history](#)

World Wide Web

From Wikipedia, the free encyclopedia
(Redirected from [World wide web](#))

"The World Wide Web" and "WWW" redirect here. For other uses, see [Web](#) and [WWW \(disambiguation\)](#). "Web surfing" redirects here. For the Web browser, see [WorldWideWeb](#).

The **World Wide Web** (commonly shortened to **the Web**) is a system of interlinked [hypertext](#) documents accessed via the [Internet](#). With a [Web browser](#), a user views [Web pages](#) that may contain [text](#), [images](#), [videos](#), and other [multimedia](#) and navigates between them using [hyperlinks](#). The World Wide Web was created in 1989 by [British](#) scientist [Sir Tim Berners-Lee](#), working at the [European Organization for Nuclear Research](#) (CERN) in [Geneva, Switzerland](#), and released in 1992. Since then, Berners-Lee has played an active role in guiding the development of Web standards (such as the [markup languages](#) in which Web pages are composed), and in recent years has advocated his vision of a [Semantic Web](#).

WIKIPEDIA
The Free Encyclopedia

navigation

- [Main page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)

search

- Used as a noun
- Actions on it – search, navigate, put, etc.

“What is ‘The Web?’” (2/2)



- A distributed document delivery system implemented using application-level protocols on the Internet
- A tool for collaborative writing and community building
- A framework of protocols that support e-commerce
- A network of co-operating computers interoperating using HTTP and related protocols to form a ‘subnet’ of the Internet
- A large, cyclical, directed graph made up of Web pages and links

Perspectives of 'The Web'



- Computer science perspective - infrastructure and intelligent systems
- Information science and knowledge management perspectives - data, information, knowledge, wisdom hierarchy
- Social intelligence perspectives - connectivity, social network intelligence
- Application perspectives - E-commerce, etc.

Perspectives of 'Science'



- Physical/biological science perspectives -analytic disciplines that aim to find laws/processes that generate or explain observed phenomena
- Social science perspective – scholarly or scientific disciplines that deal with the study of human society and of individual relationships in and to society
- Computer science perspective - a basically synthetic discipline that creates mechanisms (e.g., formalisms, algorithms, etc.) in order to support particular desired behavior

What Science Explains the Web?



- **Given**
 - Neither the Web nor the world is static
 - The Web evolves in response to various pressures from
 - ✦ Science
 - ✦ Commerce
 - ✦ The public
 - ✦ Politics
 - ✦ Etc., etc.

Web Science

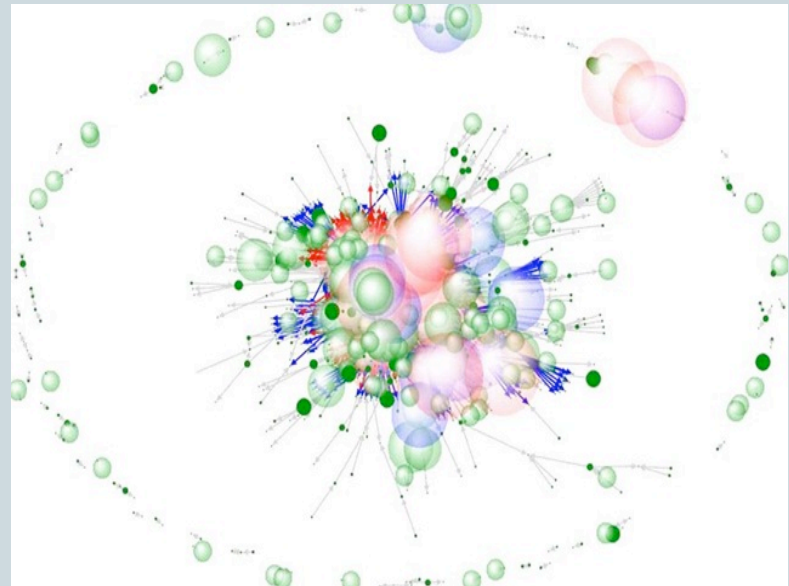
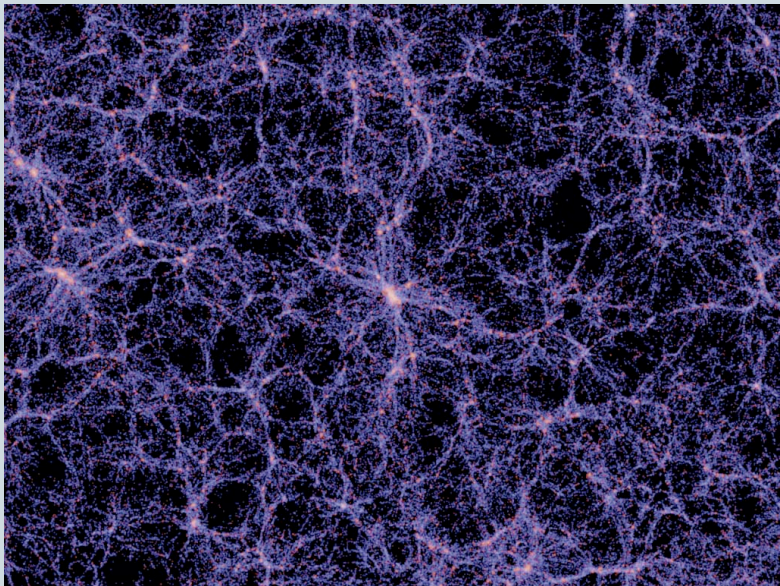


- The Web is a new technical and social phenomenon and a growing organism
- The Web needs to be studied *in situ* and understood and it needs to be engineered
- **Web Science** is a **new** field of science that involves a multi-disciplinary study and inquiry for the understanding of the Web and its relationships to us

I Would Prefer 'Web Cosmology'



- The scientific study of the origin, evolution, and structure of the universe (or Web)
- A specific theory or model of the origin and evolution of the universe (or Web)

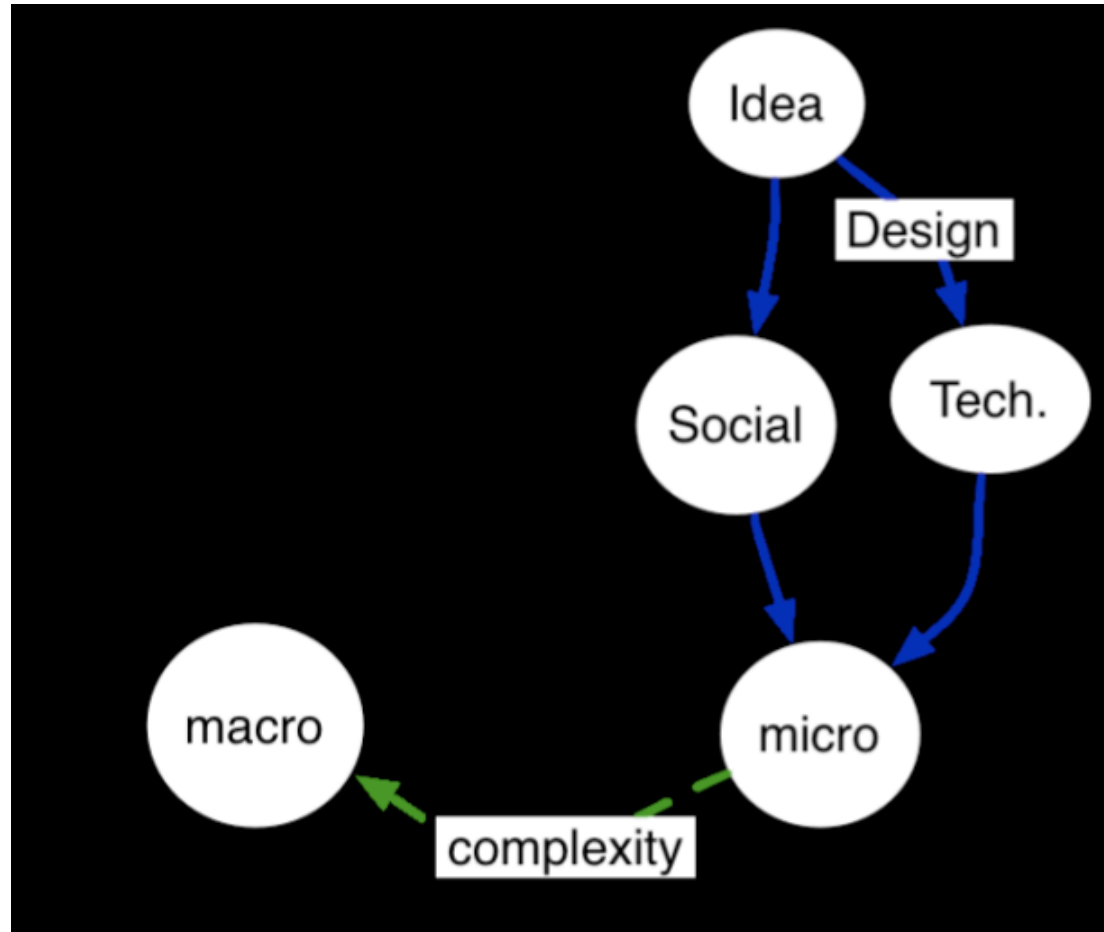


It's an Issue of Scale (1/2)



- At the **micro scale**, the Web is an infrastructure of artificial languages and protocols; it is a piece of engineering
- But the linking philosophy that governs the Web results in emergent properties (complexity) that occurs at at a **macro scale**
- The Web's use becomes a part of a **wider system of human interaction** governed by conventions and laws

It's an Issue of Scale (2/2)



(From Tim B-L)

Why Web Science?



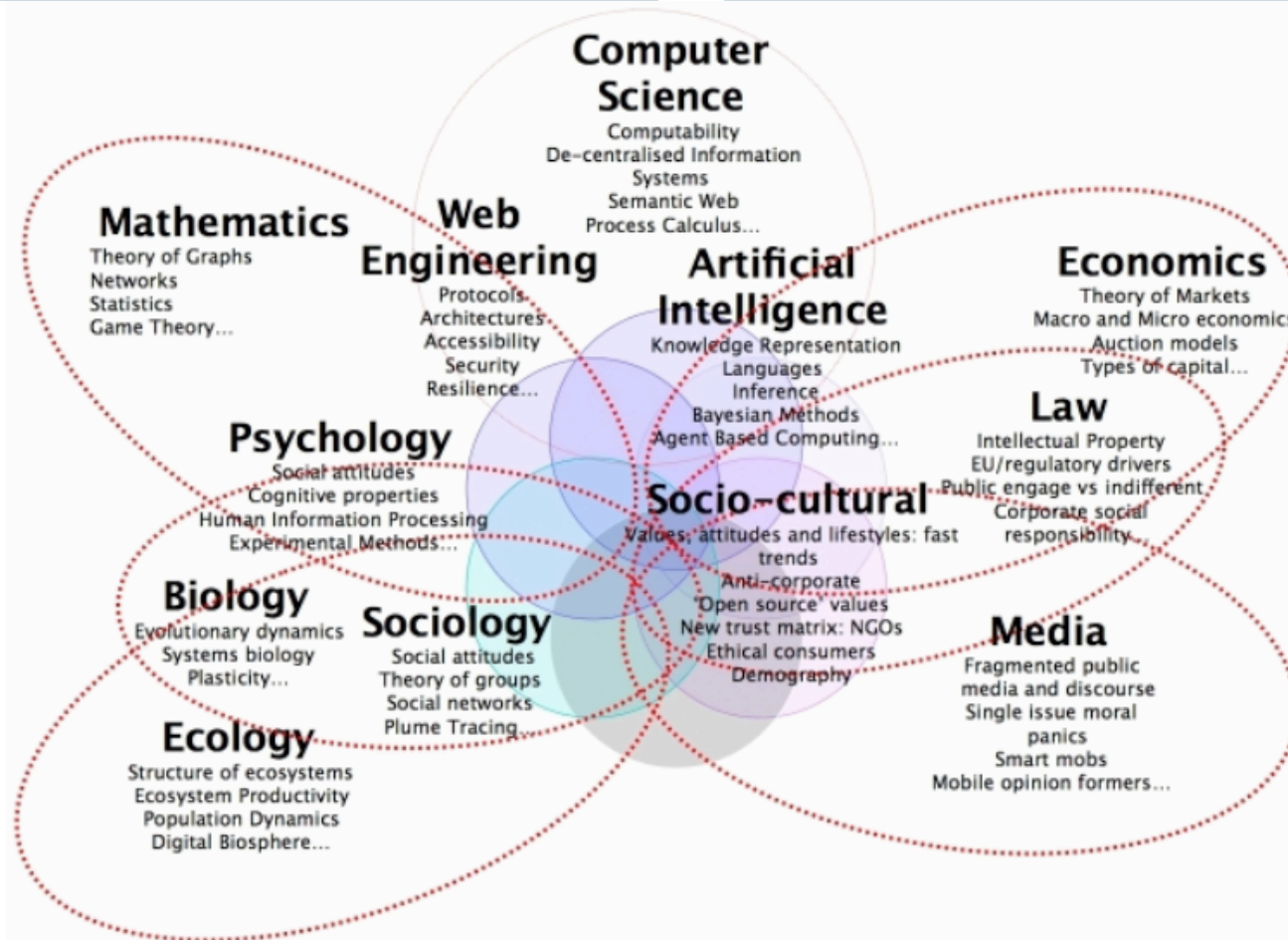
- Dynamics and evolution
- The “deep (or dark) Web”
- Sampling, lack of complete enumeration
- Scale (e.g., What is the percentage of Web pages updated daily?)
- Search (e.g., “What percentage of Web pages are indexed by search engines?”)
- Web topology
- Artifacts of social interactions (blogs, etc.), Web sociology

What Could Scientific Theories for the Web Look Like?



- **Some simple examples:**
 - Every page on the Web can be reached by following less than 10 links
 - The average number of words per search query is more than 3
 - A Wikipedia page contains, on average, 0.03 false facts
 - The Web is a “scale-free” graph
- **Can these statements be easily validated? Are they good theories? What constitutes good theories about the Web?**

Intersection of Disciplines



Roots of Web Science



- Web Ecology (Bernardo Huberman) – “The Web is an enormous opportunity to learn in the massive level things that sociologists and economists only dreamed in fantasy.”
- Web Engineering – “...covers the realization of solutions within the Web, its applications and its advancement, in particular its approaches, methods, models, principles and tools, which are based on the information and communication technologies of the Internet” (ISWE)

A Case For Web Science



- How can we understand?
 - The “dot-com” bust (relied on old software business models?)
 - Phishing, cross-site scripting (how did Web naiveté change?)
 - Etc., etc.
- Why does a system like Wikipedia “fly in the face of reason?”
- What is the appeal of systems like Facebook, MySpace?
- How can we address?
 - Legal/ethical issues
 - Internationalization
 - Trust
- How will/can the Web affect the way we “do” science, education, governance, communication, etc.?
- How will a “Web of objects” operate?
- These are **not** technical questions

The Goals of Web Science



- To understand what the Web is
- To engineer the Web's future
- To ensure the Web's social benefit

Computer Science vs. Web Science (or Why Web Science is Not Computer Science)



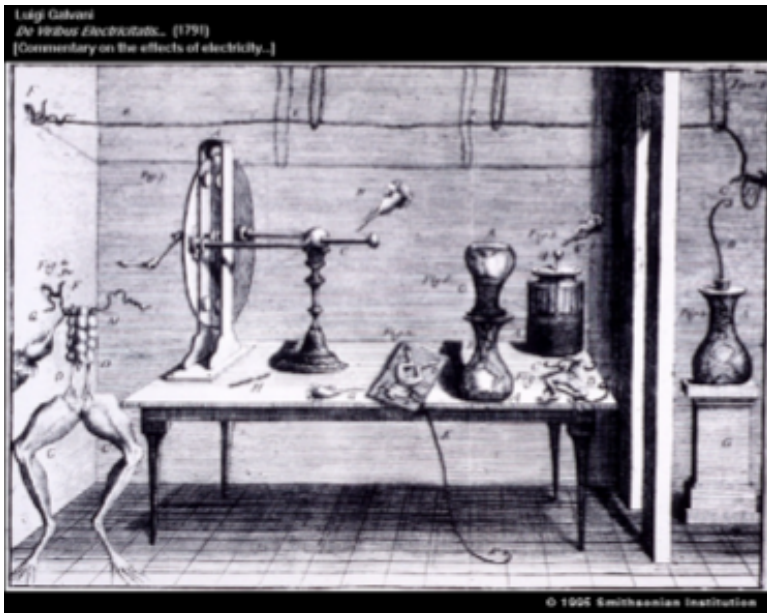
<u>Metrics</u>	
Moore's Law	Page views
Order (n) algorithm analysis	Unique visitors/month
Gigabytes	# of songs/videos
<u>Topics</u>	
Computer networks	Social networks
Packet switching	VOIP, music sharing
Information	Relationships
Programming languages	Wikis, blogs, tagging
DBs, OSs, compilers	E-*
3D graphics, rendering, etc.	Creating/sharing multimedia
<u>Focus</u>	
Technology	Applications
Computers	Users
HPC	Mobile devices, clusters
Proficient programmers	Universal accessibility

(Adapted from Ben Schneiderman)

Food For Thought



Electricity : 1800



Electricity Now

$$\oint \vec{E} \cdot d\vec{A} = \frac{q}{\epsilon_0}$$

$$\oint \vec{B} \cdot d\vec{A} = 0$$

$$\oint \vec{E} \cdot d\vec{s} = -\frac{d\Phi_B}{dt}$$

$$\oint \vec{B} \cdot d\vec{s} = \mu_0 i + \frac{1}{c^2} \frac{\partial}{\partial t} \int \vec{E} \cdot d\vec{A}$$

What are the analogies for Web Science and Design?

Breaking New Ground Together



- **Unexplored territory in Web science and engineering**
 - Broad scope for research agenda
 - New relationships among theoreticians, experimentalists, and systems and applications builders
 - New relationships with social science, law, economics, psychology, etc.

Challenges



- **Web Science suggests that**
 - We can use non-technical concepts to understand the complexity of our Web applications so that we can engineer them to have new and predictable behaviors
 - We can better understand the impact of Web technology in all areas of communication and social interaction
 - We may be able to reliably predict the future evolution of the Web



Thank You!

Questions?/Comments?

bebo@slac.stanford.edu