What Have We Learned From Web Science? *A Status Report*





HKU ECOM-ICOM EXPERT ADDRESS 26 SEPTEMBER 2012 "It is a privilege to be present at the birth of a new science"

Agenda

- Overview of Web Science
- Efforts toward a Web Science Curriculum
- Examples of Web Science research
- Conclusions (and a call to action!)
- References

Web Science Trust (WST)

- http://webscience.org/
- 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."

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

What Disciplines Can the Web Give Insight Into?

- (Besides Computer Science and Engineering too obvious)
- Sociology
- Psychology
- Mathematics
- Biology
- Etc.,etc.

Web Science (1/2)

- **Given**: The Web is a new technical and social phenomenon and a growing <u>organism</u>
- **Asserted**: The Web needs to be studied <u>in situ</u> and understood and it needs to be engineered
- Not to be confused with Computer Science or "doing science" on the Web
- <u>Web Science</u> is a <u>new</u> field of science that involves a multi-disciplinary study and inquiry for the understanding of the Web and its relationships to us
- Brought formally to academic and public attention in 2006

Web Science (2/2)

- A new science that focuses specifically on applications in multiple disciplines
- Its principal questions reflect a fundamental curiosity about the Web tempered by the awareness that a question's importance scales with respect to its relevance to a societal imperative
- Conference series began in 2009



No data available

Web Science is a Result of Scale (1/2)

- At the <u>micro scale</u>, 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 <u>macro</u> <u>scale</u>
- The Web's use becomes a part of a <u>wider system of</u> <u>human interaction</u> governed by conventions and laws (?)



Why Web Science?

- Dynamics and evolution
- The "deep (or dark) Web"
- Sampling, lack of complete enumeration
- Search (e.g., "How can the Web's data best be found and utilized?")
- Artifacts of social interactions (social networks, Web sociology)
- Web topology
- Is the Web being used to its fullest extent?

A Case For Web Science (1/2)

• How can we understand?

- The impact of the Web on business models
- Web-unique legal and societal issues
- How people use the Web
- Etc., etc.

Web Science is "bi-directional"

- We are not just learning about the Web
- We are using the Web and its data, topology, operation, etc. (the Web ecosystem) as a laboratory for many other disciplines

A Case For Web Science (2/2)

- What is the appeal of systems like Facebook, (~900 million registered users)?
- How can we address?
 - Internationalization
 - o Trust
- How will/can the Web affect the way we "do" science, education, governance, communication, etc.?
- "Web of Objects" "Big Data", etc.
- Not all of these are technical questions or have technical answers/solutions



Intersection of Disciplines (2/2)

- Each of these disciplines are "cells"
 - That have difficulty communicating
 - That have difficulty "exchanging DNA"

• Will continue to "affect" and "be affected by" the Web



The Goals of Web Science

- To understand what the Web is
- To engineer the Web's future
- To ensure the Web's social benefit
- To effectively use the Web as a tool

A New Kind of Science

"Science should no longer seek theories that scientists can understand, because the digital cloud will understand them better anyway" ---Chris Anderson, WIRED

Breaking New Ground

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.
- How should this be taught? How do we go about teaching "Webiety?"



(University of Southampton)

Example – SH PhD Topics

- "Principles of Governance for Open Scientific and Government Law" (law,CS,linguistics)
- "Social Web, Social Behavior & Complex Adaptive Networks" (CS,psychology)
- "Large Scale Data-Stream Mining of Distributed Network Systems" (CS,social statistics)
- "Cybercrime Target Hardening" (CS,criminology)
- "Intermediation with e-Tailing 2C Supply Chain" (management,CS)
- "How Does the Web Grow? Analyzing the Emergence of Linked Data" (sociology,CS)



Green: large participation Yellow: moderate participation Red: low participation

Web Science and the Social Sciences

- "Web Science embraces the study of the Web as a vast information network of people and communities. It also includes the study of people and communities using the digital records of user activity mediated by the Web. An understanding of human behavior and social interaction can contribute to our understanding of the Web, and data obtained from the Web can contribute to our understanding of human behavior and social interaction." [ACM Web Science conference site]
- "Studying the online world to understand the offline world."

Two Examples

Sociograms & "Small-World"

- A graphic representation of the social links that a person has
- Can include social relations, channels of communication, social influence, etc.
- o Jacob Moreno, Stanley Milgram
- Can the Web provide a viable laboratory?

Augmented intelligence through crowdsourcing

- Not to confused with Artificial Intelligence
- Intelligence amplification (Engelbart)
- Can we measure the wisdom/intelligence of the "crowd?"

Example 1 - Sociograms







- 99.6% of all pairs of users are connected by paths with 5 degrees (6 hops)
- 92% of users are connected by only 4 degrees (5 hops)
- As Facebook has grown, average number of hops has reduced (2008 – 5.28; 2011 – 4.74)



(Laboratory for Web Algorithmics, Università degli Studi di Milano)

Does the "Offline World" Reflect the "Online World?"



25% of all Facebook users make up 75% of all "friendship links"

Example 2 – "Intelligence of the Crowd"

- Microsoft researchers used Amazon Mechanical Turk workers to answer questions drawn from a standard IQ test to establish a proxy for the intellectual potential of the "crowd" i.e., a crowd IQ score
- Also studied factors influencing crowd IQ score such as
 - Payment
 - Number of workers per task
 - Worker reputation
 - Rejection rules

Choose a correct answer to this computer generated reasoning problem. Guidelines:

- Some questions might be very difficult.
- There is only one correct response.
- NO REJECTIONS for incorrect answers...
- ... BUT please DO NOT SPAM it is very easy to recognize spammers and we will block and reject them.



Sample Results

- Crowd IQ as a function of number of workers per Human Intelligence Task (HIT)
- All payment conditions included
- Reputation > 98%



Conclusions/Status of Web Science

- Interest in the field is growing albeit at a slow pace
- Major institutions are establishing Web Science curricula – where do the graduates of these curricula go?
- Very interesting work has been done in some of the related disciplines
- Web Science initiatives need to be started in a variety of disciplines

Challenges

• Web Science suggests that

- The Web can effectively be used as a laboratory for a wide variety of disciplines
- We should be able to use non-technology-specific concepts to explain the complexity of Web principles in order to make them accessible to the widest possible audience
- We should be able to engineer the Web and its data for new and innovative purposes
- We can hopefully 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

WSTNet (1/2)

- The WST has established a new international network bringing together world-class research laboratories to support the WS research and education program
- WSTNet combines some of the world's leading academic researchers in WS, with new academic programs that will enhance the already growing influence of WS. Through a number of specific agreements and commitments with the WST, which will manage WSTNet, the member Labs will provide valuable support for the ongoing development of WS.
- Contributions from the Labs will include the organization and hosting of summer schools, workshops and meetings, including the WebSci conference series. The WSTNet labs will also identify new opportunities for additional events and fundraising, all as part of the ongoing development of WS.

WSTNet (2/2)

- USC (USA)
- VU Amsterdam (Netherlands)
- Koblenz (Germany)
- Northwestern (USA)
- RPI (USA)
- L3S Hannover (Germany)
- PUC (Brazil)

- MIT (USA)
- DERI (Ireland)
- Oxford (UK)
- Southampton (UK)
- Tsinghua-Southampton (China)
- KAIST (South Korea)
- IU (USA)
- YOU

References

- The Web Science Trust (webscience.org)
- WebSci'13 May 2-4 2013 Paris, France
- WWW2013 May 13-17 2013 Rio de Janeiro, Brazil

Thank You!

Questions?/Comments?

bebo@slac.stanford.edu