Observatories and data analytics for Web Science Research

Professor Dame Wendy Hall
University of Southampton, UK
25th August 2015

@DameWendyDBE
The Web – most successful information architecture in history
The Age of Data

• A Web of **linked data** was always part of Tim’s original vision

• Machines can process and interpret linked data to make inferences about that data leading to a more intelligent (semantic) Web

• **Open data** leads to greater transparency, efficiency, and economic and social value, as demonstrated by the UK’s Open Data Institute. **Linked open data** is even more powerful

• **Big data** – the Web has enabled the generation of lots of data that we are hungry to analyze and share
Why Web Science?
The Web – we need to understand it

• Web architecture is simple set of protocols

• These give rise to complex macro phenomena

• Need systems oriented view of the Web and its ecosystem – Web Science

• One that acknowledges social and technical components
“Web Science represents a pretty big next step in the evolution of information. This kind of research is likely to have a lot of influence on the next generation of researchers, scientists and, most importantly, the next generation of entrepreneurs who will build new companies from this.”

Dr Eric Schmidt, CEO, Google Inc.
Not the union of the disciplines

But more than their intersection

Web Science and additionality
www.webscience.org

Exists to promote and help coordinate

– Research
  • Research Agenda
  • WSTNet Labs

– Education
  • Curriculum Development
  • Summer Schools

– Thought Leadership
  • Workshops
  • Dissemination

info@webscience.org
The Web Science Network of Laboratories (WSTNet) combines some of the world’s leading academic researchers in Web Science, with new academic programmes that will enhance the already growing influence of Web Science. The member Labs will provide valuable support for the ongoing development of Web Science. There are now 15 WSTNet labs:

Southampton, UK
MIT, USA
North-Western, USA
Tsinghua, China
DERI, Galway, Ireland
KAIST, Korea
L3S, Hannover, Germany

Oxford, UK
RPI, USA
Anaheim School of Communication, USC, USA
VU, Amsterdam, The Netherlands
Koblenz, Germany
Rio, Brazil
Indianna, USA
NUS, Singapore
The Web as a Social Machine
Web Science is the theory and practice of social machines?

“Real life must be full of all kinds of social constraints – the very processes from which society arises. Computers help if we can use them to create abstract social machines on the Web, processes in which people do the creativity and the machine does the administration. The stage is set for an evolutionary growth of new social engines. The ability to create new forms of social process would be given to the world at large and development would be rapid”

Tim Berners-Lee and Mark Fischetti, Weaving the Web, 1999
Examples of social machines

- The Web, Google, Facebook, Twitter, Wikipedia ……..
- Trip Advisor
- Zooinverse
- Ushahdi – open source project which allows users to crowd source crisis information to be sent via mobile
- The OpenStreet Map of Haiti created after the earthquake
- The list goes on … Amazon, e-Bay, YouTube ….design your own
Social Machines in Context

- More machines
- More people

- Big Data
- Big Compute

- Conventional Computation

- Social Machines
- Social Networking
Social Machines are NOT Turing Machines

- they do contain conventional algorithmic components but much else is different

- a social machine will start with an incomplete specification that grows and evolves to cover more of the problem via interaction

- a social machine achieves participation through local incentives which become reinforced as the...

- incentive for an individual to supply data to the algorithm increases as more individuals participate

- a social machine has a notion of completeness that is a social rather than mathematical issue

- a social machine will not usually have a notion of the correct output or termination... rather it runs continuously
Now I understand ..... 

- If Web Science is the theory and practice of social machines
- Then Computer Science is the theory and practice of Turing machines
- What about a Turing test for social machines?
- How do we study social machines?
The Web Observatory

Tiropanis, Hall, Shadbolt, DeRoure, Contractor & Hendler
“The Web Science Observatory”
IEEE Intelligent Systems, May 2013
Web Science across continents

- Astronomers obtain a very high resolution picture of the sky from small telescopes a long distance apart.
- Many labs, contributing across the globe, help build an accurate picture of human activity at planetary scale.
  - transcending parochial social, political, economic, legal interpretations
Understanding climate change

Think of the Web Observatory as a tool for a kind of SOCIAL climatology, e.g. Like studying climate understanding society on the Web and via the Web requires long term study of complex interacting patterns of data and usage. This data is gathered from many sources, brought together and analysed to help us understand our complex and ever-changing Web climate.
Web Observatory: Global partnerships

• Partners contribute their insight and experience, and benefit from the network and business intelligence insights

• Observatory events are hosted bringing together thought leaders to learn from each other

• Data sets, **open or closed**, can be shared under t’s and c’s

• Analytics and tool sets can be contributed

• Joint research and projects can be agreed

• This enables longitudinal research
understanding web evolution:
- observation
- experimentation
Levels of Sharing

- Professional & Organisational Datasets
- Shareable Datasets
- Open Datasets
- Shared Analytics and Visualisation Tools

Thanassis Tiropanis – University of Southampton
We are building a social machine to observe social machines
The Southampton Web Observatory
The case for interoperability (Linked Data)

Web Observatory Portals

- Community engagement tools
- Sharing tools
- Authoring & publication tools

Sharing protocols

Discovery/Advertisement schemas and protocols

Datasets

- Datasets catalogue
- Triple stores
  - Dataset
  - Dataset

Enrichment
- Management
- Harvesting
- Wikipedia
- Microblogging
- Streaming media

Analytic tools

- SPARQL (Distributed)
  - Analytic tools catalogue
    - Web hosting
      - Trends
      - Simulations
      - Visualisations

http://webobservatory.soton.ac.uk

Welcome to the Southampton University Web Observatory

Part of the global web observatory

Datasets
A page containing information regarding various datasets gathered
View details »

Visualisations
A page containing various data visualisations
View details »
Observatory in action

- Developing a live environment to observe and analyze in real-time

Weibo: Anti-Corruption messages and network.

Chinese ‘Salt Crisis’: Visuals of Humour
Observations: Twitter and Wikipedia

- Observing the effects of real-world events across multiple sources

  - Combining multiple streams of real-time information to better understand and more accurately predict events
The Web of Observatories
How do we catalogue Observatories and content?
Getting started

https://www.w3.org/wiki/WebSchemas/SchemaDotOrgProposals

https://www.w3.org/wiki/WebSchemas/WebObsSchema
List of Observatories

Observatories may sit in many different places gathering data about different topics. A powerful effect emerges when one Observatory can discover another and share its data and tools.

http://webscience.org/web-observatory/
Welcome to IIIT-Bangalore Web Observatory

Part of the global web observatory

Datasets
A page containing information regarding various datasets gathered
View details »

Visualisations
A page containing various data visualisations
View details »
Welcome to the UniSA Web Observatory

Part of the global web observatory network

Datasets
A page containing information regarding various datasets gathered

View details »

Visualisations
A page containing various data visualisations

View details »

Links
Governance in a Social Machine Ecosystem - Observing the Web as a way of making sense of a world of "wicked" problems

Government as a Social Machine report 1: The implications of government as a 'social machine' for making and implementing market-based policy

Government as a Social Machine report 2: The Machines
The ambition is to map the digital universe
UI Summer School Projects

Data from UoS and UI Web Observatory

Predicting Election Popularity based on interactions and language

Detecting structures of political parties based on geographic location. Using Sentiment as a way to understand structure

Creating new metrics to measure the popularity and influence of political parties. What makes a political-social influencer

Designing smart city scenarios to improve citizen awareness and day-to-day movement

Using Smart City data to model events and their corresponding social sentiment
Twitter Conversations during US Elections 2012

Twitter Conversations during UK Elections 2015
Traffic in blue dots
Parking in orange diamonds
What next for the Web?
The Future as the Web turns 25

- Amazing technical developments ahead but also major challenges – internet governance, net neutrality, cybersecurity, privacy, trust, ....

- Who has the right to do what with our data – fallout from the Snowden affair and development of personal data stores

- A human rights charter for the internet?

- We need evidence to help drive policy and provide business intelligence. This means an observatory approach to data analytics for the Web
“Power relies on the control of communication. Digitization of everything implies digital surveillance can be comprehensive and this is unprecedented. Under these conditions democracy is threatened”

Manuel Castells, Cambridge, UK, March 2015
eGovernment and Digital Democracy

• Redefining the relationship between government and citizen

• What is the impact of social media on the election process?

• Will it change the way we elect our governments and make our laws?

• To answer these questions we need an interdisciplinary, evidence based, observatory approach
Social Machines in Context

More machines

Big Data
Big Compute

Conventional Computation

Social Networking

More people
Thank you

Questions?