Towards a Smarter Web:

Professor Dame Wendy Hall
19 July 2011
My Inspiration
“As we may think”
Vannevar Bush
Atlantic Monthly
July 1945
Everything is deeply intertwined.

Ted Nelson and Doug Engelbart

Augmenting human intellect
Hypertext comes of age: 1987

Hyperland
What do India, the Earl Mountbatten of Burma, the University of Southampton and my research career have in common?
The Mountbatten archive moved to Southampton in 1987.
Microcosm: Mountbatten archive application

The 2nd miscellaneous meeting of His Excellency the Viceroy was held at Viceroy's House, New Delhi, at 11.15 hours on Wednesday, 2nd April 1947.

The item discussed was:-

The Indian National Army

Present

The Viscount Mountbatten of Burma, Viceroy of India.
The Hon'ble Pandit Nehru, Member for External Affairs and Commonwealth Relations.

* The Hon'ble Patel, Member for Home Affairs, Information and Broadcast.
The Hon'ble Liaquat Ali Khan, Member for Finance.
The Hon'ble Sardar Baldev Singh, Member for Defence.

Field Marshal Sir Claude Auchinleck, C in C [CINC] India.
Associative Linking

“An associative link is the instantiation of a semantic relationship between information elements. The semantic relationship can be something as simple as ‘definition of’ or something far more complex.”

Lowe & Hall, Hypermedia and the Web: An Engineering Approach, Wiley, 1999
Linkbases in Microcosm

Note the direction of this arrow!

Link database

Separable hyperstructure

Documents
Links in Microcosm

- source, destination, description
- source: object | concept | context
- We generated links based on metadata description of documents in docuverse and “it all falls out”
ACM Hypertext’91
The Web – most successful information architecture in history
Lessons learnt

• Big is beautiful: the network is everything
• Scruffy works: let the links fail to make it scale
• Democracy rules: open, free and universal

• But we lost (for a time) conceptual and contextual linking, and link descriptions – the Web is a strangely linkless world
• Missing links – search engines fill the gap
Google

Search 1,387,529,000 web pages

Google Search   I'm Feeling Lucky

Google Web Directory  Google Groups
the web organized by topic  usenet discussion forum

Do-it-yourself keyword advertising. Google AdWords works.

Chad, one of our Googlers, is biking across America!

Cool Jobs - Add Google to Your Site - Advertise with Us - Google in Your Language - All About Google

Make Google Your Homepage!

©2001 Google
The Proxy DLS

The DLS Agent is an HTTP proxy that inserts links on-the-fly by querying the link service.
Link Service Evolution at Southampton

DLS
OHP
FOHM
Auld Linky
XML  XLink  XPointer

- Open Hypermedia Systems community influence on the Web?

- What happened to contextual and conceptual linking?
COHSE (1999)

- Improving the quality, consistency and breadth of linking of Web documents using ontologies.
The Semantic Web

Semantic Web LayerCake (Berners-Lee, 99; Swartz-Hendler, 2001)
The Semantic Web

Semantic Web LayerCake (Berners-Lee, 99; Swartz-Hendler, 2001)

- RDF triples: Subject, Predicate, Object
The Web – exponential growth and impact
Web 2.0

- Wiki’s
- Blogs
- Flickr
- YouTube
- MySpace
- Facebook
- Second Life
- Twitter
Timeline of Twitter Growth

Winkin Huang Yonhjian PhD student
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.
webscience.org

• Exists to promote and help coordinate

  – Research
    • Research Agenda
    • WSTNet Labs
    • ACM Web Science 2011 Koblenz 15-17 June

  – Education
    • Curriculum Development
    • Summer Schools – DERI 6-13 July 2011

  – Thought Leadership
    • Workshops
    • Dissemination

  – Your help and support
    • info@webscience.org
Web Science and additionality

Not the union of the disciplines

But more than their intersection
The Web – emergent shape and structure

Scale-free
- Some nodes are of high degree most are low degree

Power laws
- The degree distribution follows a power law, with an exponent $\beta > 2$.

Small worlds
- The average distance (or diameter) is much smaller than the order of the graph.

Hubs and authorities
- The number of distinct bipartite cliques or cores is large when compared to a random graph with the same number of nodes and edges.
The Web – exploiting shape and structure

\[
I^0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n \\
I^1 = SI^0 = c_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n \\
I^2 = SI^1 = c_1 v_1 + c_2 \lambda_2^2 v_2 + \ldots + c_n \lambda_n^2 v_n \\
\vdots \\
I^k = SI^{k-1} = c_1 v_1 + c_2 \lambda_2^k v_2 + \ldots + c_n \lambda_n^k v_n \\
GI^k = \alpha H I^k + \alpha A I^k + \frac{1-\alpha}{n} I^k
\]
The Web – a socio-technical system

Web Science is the theory and practice of social machines
The Web – emergent networks

Courtesy Mathew Hurst
The Web – emergent problem solving
The Web – emergent problem solving

Crowd sourcing

Words Digitized Per Day

Scanned book

Word that OCR can’t read

Randomly distorted image of word

upon

between

Type the two words:
The Web – emergent data
The Semantic Web
A Web of Linked Data
"Why the Semantic Web will Never Work"
(note the quote marks!)

Jim Hendler
RPI
http://www.cs.rpi.edu/~hendler
@jahendler

10 years on, ESWC 2011
(sorry, not in rhyme)
Content, Emergence and Unanticipated Reuse

The four micro principles of the Semantic Web

1. All entities of interest, such as information resources, real-world objects, and vocabulary terms should be identified by URI references.

2. URI references should be dereferenceable, meaning that an application can look up a URI over the HTTP protocol and retrieve RDF data about the identified resource.

3. Data should be provided using the RDF/XML syntax.

4. Data should be interlinked with other data.
Linked Data on the Web: May 2007

- 500 Million RDF Triples
- 120,000 RDF links between data sets
Linked Data on the Web: April 2008

23 billion RDF Triples
3 million RDF links between data sets
LOD Datasets on the Web: March 2009

4.5 billion triples
180 million data links

As of March 2009
Tipping points?
January 1, 2009

"Openness will strengthen our democracy and promote efficiency and effectiveness in Government."

--- President Obama

May 21, 2009

data.gov online
57 Data Sets

December 8, 2009

“Open Government Directive” released
~2000 Data Sets

May 21, 2010

data.gov relaunch with semantic web featured

June 30, 2009

Putting Govt Data online- Data.gov.uk beta

January 19, 2010

~3000 Data Sets

June 2010

~4500 Data Sets

2009

2005

The Web – open government data
Open Government Data
Open Local Authority Data
Open City Data

Welcome to the London Datastore

The London Datastore has been created by the Greater London Authority (GLA) as an innovation towards freeing London's data. We want citizens to be able to access the data that the GLA and other public sector organisations hold, and to use that data however they see fit - free of charge. The GLA is committed to informing and engaging other public sector organisations into releasing their data too.

Relating data though is just half the battle. Raw data often doesn't tell you anything until it has been presented in a meaningful way. We want to encourage the masses of technical talent that we have in London to transform raw data into tools that citizens can actually find useful.

Join the conversation on Google+, Twitter or email and give us your thoughts and feedback - we want this to be work for you too.

Browse Catalog

Please note that data feeds contain the latest update from the source database (which might include records from several days ago and not today's date).

Displaying: All categories

<table>
<thead>
<tr>
<th>Category: All categories</th>
<th>Source</th>
<th>Metadata</th>
<th>XML</th>
<th>Text/CSV</th>
<th>Atom (GeosMash support)</th>
<th>KML/ESRI Shapefile</th>
<th>Maps</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCTO</td>
<td>District Payments To All Vendors, by Census Tract, FY10, Q1</td>
<td>District Payments To All Vendors, by Census Tract, FY10, Q1</td>
<td>District Payments To All Vendors, by Census Tract, FY10, Q1</td>
<td>08/04/2010</td>
<td>Provides information regarding payments made to vendors in FY10, Q1 for goods or services provided to the District.</td>
<td>Provides information regarding payments made to vendors in FY10, Q2 for goods or services provided to the District.</td>
<td>Provides information regarding payments made to vendors in FY10, Q3 for goods or services provided to the District.</td>
<td>Provides information regarding payments made to vendors in FY10, Q4 for goods or services provided to the District.</td>
</tr>
</tbody>
</table>
Why?

- Transparency
- Accountability
- Engagement
- Improve Public Services
- Efficiency – better for less
- Economic and social value
ASBOs, Dentists and Tubes
UK OGD – Spending…
Most recently....
The road to stardom

| ★  | Put your data on the Web (any format) |
| ★★ | Make it available as structured data |
| ★★★ | Use open, standard formats |
| ★★★★ | Use URLs to identify things |
| ★★★★★ | Link your data to other people’s data |
University of Southampton Open Data

The University of Southampton provides open access to some of our administrative data.

We believe that this will be of benefit to our own members and visitors, and increase the transparency of our operations.

Featured App:

Open Data Map

This tool is under development by Postgraduate students, it's a work in progress and so may break now and then, but it looks amazing. You can search for services on and near our campuses!

Linked Open Data

The executive summary: There's data we have which isn't in any way confidential which is of use to our members, visitors, and the public. If we make the data available in a structured way with a license which allows reuse then our members, or anyone else, can build tools on top of it without needless bureaucracy. That's common sense. We call it "Open Data".

For more on Open Data and its benefits see these presentations by Southampton's Nigel Shadbolt® and Tim Berners-Lee. They helped establish data.gov.uk the UK Government's Open Data site and are members of the GovTrack and Tim Berners-Lee's Data Parl.
Welcome to the University of Southampton Catering Services website. We pride ourselves on putting our customers first.

Within **Catering Services**, meal planning is a top priority and our first class chef's have selected a varied menu, including healthier options, fast food and international trends, to hopefully suit every taste bud.

We cater for staff, students and visitors to the University and provide venues that appeal to everyone. From the Student based Piazza together with fine dining in the Blue Room, we really do have something for everyone.

Halls of residence have great restaurants providing the very best in catering. Gina and her team are on hand to help with your dietary needs and can advise you on using your 'Catered Student' cards.

Open Days through to Graduation and celebrations!
Welcome to the University of Southampton Catering Services website. We pride ourselves on putting our customers first.

Within Catering Services, meal planning is a top priority and our first class chef’s have selected a varied menu, including healthier options, fast food and international trends, to hopefully, suit every taste bud.

We cater for staff, students and visitors to the University and provide venues that appeal to everyone. From the Student based Piazza together with fine dining in the Blue Room, we really do have something for everyone.

Halls of residence have great restaurants providing the very best in catering. Gina and her team are on hand to help with your dietary needs and can advise you on using your 'Catered Student' cards.

Open Days through to Graduation and celebrations! We can look after your every need.
Open Data Service

Arlott Bar

Currently closed.

Disability Information: View Disability Report for this Location
Located within: Staff Social Centre

Monday: 09:00-18:00
Tuesday: 09:00-18:00
Wednesday: 09:00-18:00
Thursday: 09:00-18:00
Friday: 09:00-21:00

The Arlott Bar caters for staff and post graduate students at the University of Southampton, in a warm and calming environment. There is a state of the art sound and visual display facility, Sky television and a pool table.

The Arlott Bar caters for staff and post graduate students at the University of Southampton, in a warm and calming environment. There is a state of the art sound and visual display facility, Sky television and a pool table.
Open Data Service

Apps Using Our Data

These are Apps using data from the University. If you would like your App here, please let us know, also let us know if there's an App you'd like to see.

Amenity Map

Searchable map for finding buildings, amenities and bus stops in and around University sites

Type: Web
Created: 2011-03-06
Created by:
- Janutas Pattanaphanchai - maito:jp11g09@ecs.soton.ac.uk
- emax - maito:emax@ecs.soton.ac.uk
- Colin R. Williams - maito:crw104@ecs.soton.ac.uk

Uses datasets: Buildings and Places, Teaching Room Features, Local Amenities, Catering, Southampton Bus-stops, Southampton Bus-routes

Embeddable Map Widget

Embed maps of our campuses into conference, lecture pages, or blogs.

Type: Embeddable Web Widget
Created: 2011-03-06
Mobile Bus Stop Finder

Tool to discover nearby bus-stops by GPS or postcode. Then gives live times for buses at that stop.

Type: Mobile Web
Created: 2011-03-06
Publisher: University of Southampton
Created by:
  • Christopher Gutteridge - mailto:cjg@ecs.soton.ac.uk

Uses datasets: Southampton Bus-stops

Phonebook

Type: Embeddable Web Widget, Web
Created: 2011-03-04
Publisher: University of Southampton
Created by:
  • Dave Challis - mailto:dsc@ecs.soton.ac.uk

Uses datasets: Organisation, Public Phonebook

SotonBus

Type: iPhone/iPad
Created: 2011-02-16
Publisher: University of Southampton
Created by:
  • John Lasger - mailto:j1x07@ecs.soton.ac.uk
  • Ashley Browning - mailto:amb1908@ecs.soton.ac.uk
Conclusion: “Why the Semantic Web will never work”?

Tetherless World Constellation

- No reason at all
  - The Semantic Web is here, it is working, and it will continue to do so

- But, for it to move to the next level and be all that we as a community have aspired for
  - We must revisit and update the early visions for the modern web
  - We must unify the “competing” models of linked-data and machine-readable vocabularies
  - We must step up to some critical research challenges
Towards a Smarter Web: The Return of the Link

• The exciting thing about linked data is the linking
• Revisit semantic web link services
• Triple stores <=> link databases
• It’s not all about visualising huge amounts of linked data (specialised applications)
• Think about integrating the document and data webs RDF -> RDFa
• Where will schema.org take us?
• Keep the onotologies simple initially – let them evolve
• What happened to the agents!?
• The hard work is yet to come!!!
Everything really is deeply intertwined