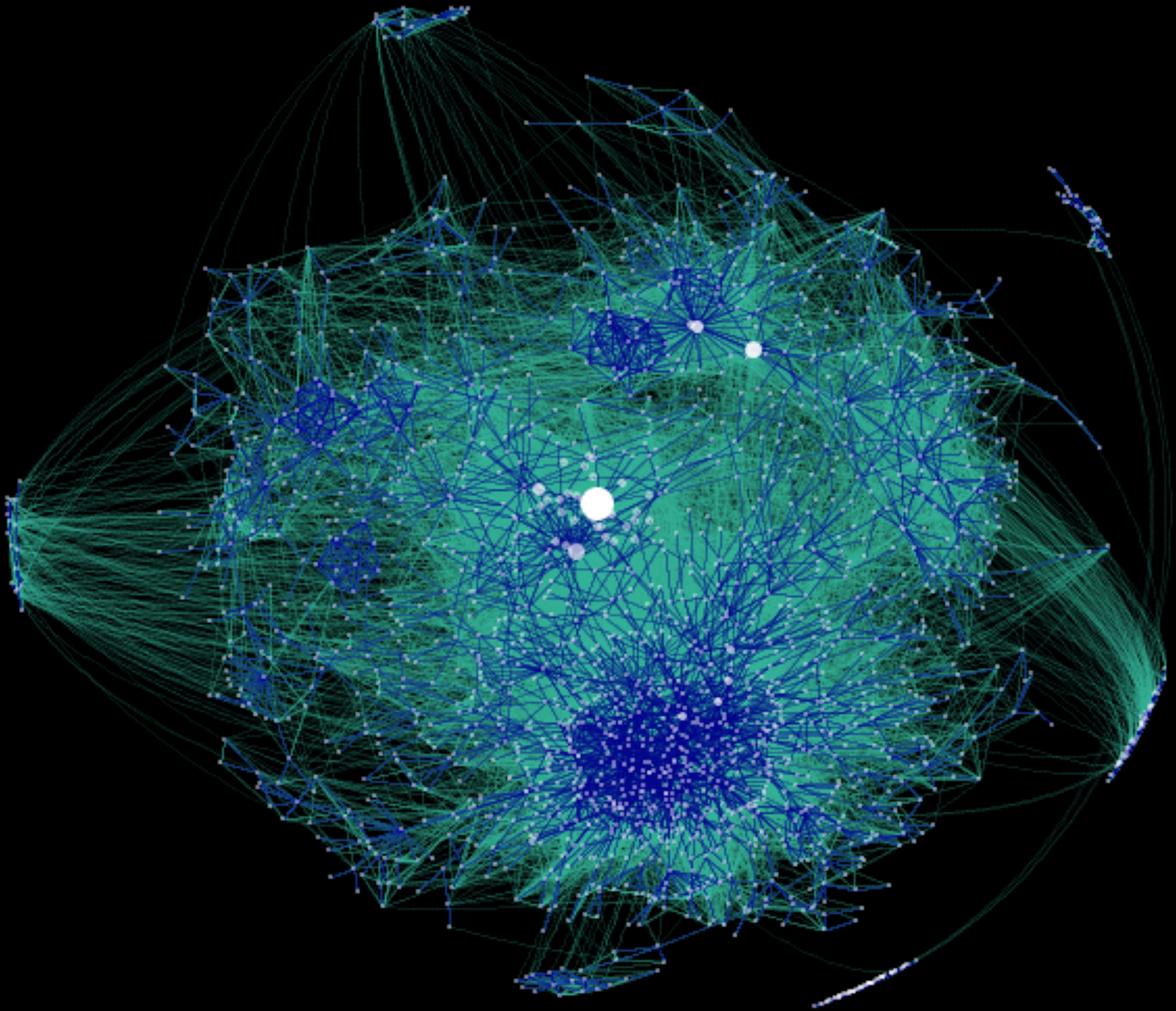


Intro to Web Science

September 19, 2013
ITWS 1100

John Erickson, Kristine Gloria, Qingpeng Zhang
Tetherless World Constellation
Rensselaer Polytechnic Institute

Agenda

1. A Science of The Web and why it matters
2. Web Architecture/Engineering the Web
3. Measuring the Web
4. The Web Science Method
5. Social Aspects of the Web
   a. Evolution of methodology
   b. Hurdles of incorporating the “social”
   c. Why humans aren’t just "nodes" in a network
6. Web and other Governance

What is Web Science?

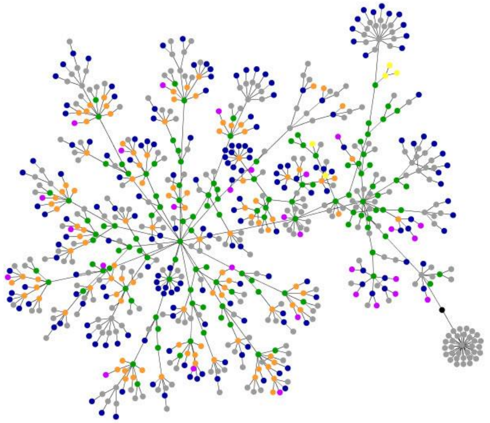
- Positions the World Wide Web as an object of scientific study unto itself
- Recognizes the Web as a transformational, disruptive technology
- Its practitioners focus on understanding the Web...
  - ...its components, facets and characteristics
- The Web Science Method: “the process of designing things in a very large space..."

What does Web Science ask?

- What processes have driven the Web’s growth, and will they persist?
- How does large-scale structure emerge from a simple set of protocols?
- How does the Web function as a socio-technical system?
- What drives the viral uptake of certain Web phenomena?

Bottom line: *What might fragment the Web?*

What is the Web?



- "The Web is not a thing..."
- **Continuously changing** due to coordinated *and* conflicting processes
- An **evolving large-scale structure** dependant on static *and* emerging protocols
- A **socio-technical system** that reflects *and* obfuscates social and technical structures
- Always goes where we *allow* it to go...but seldom where we *want* or *expect* it to go!

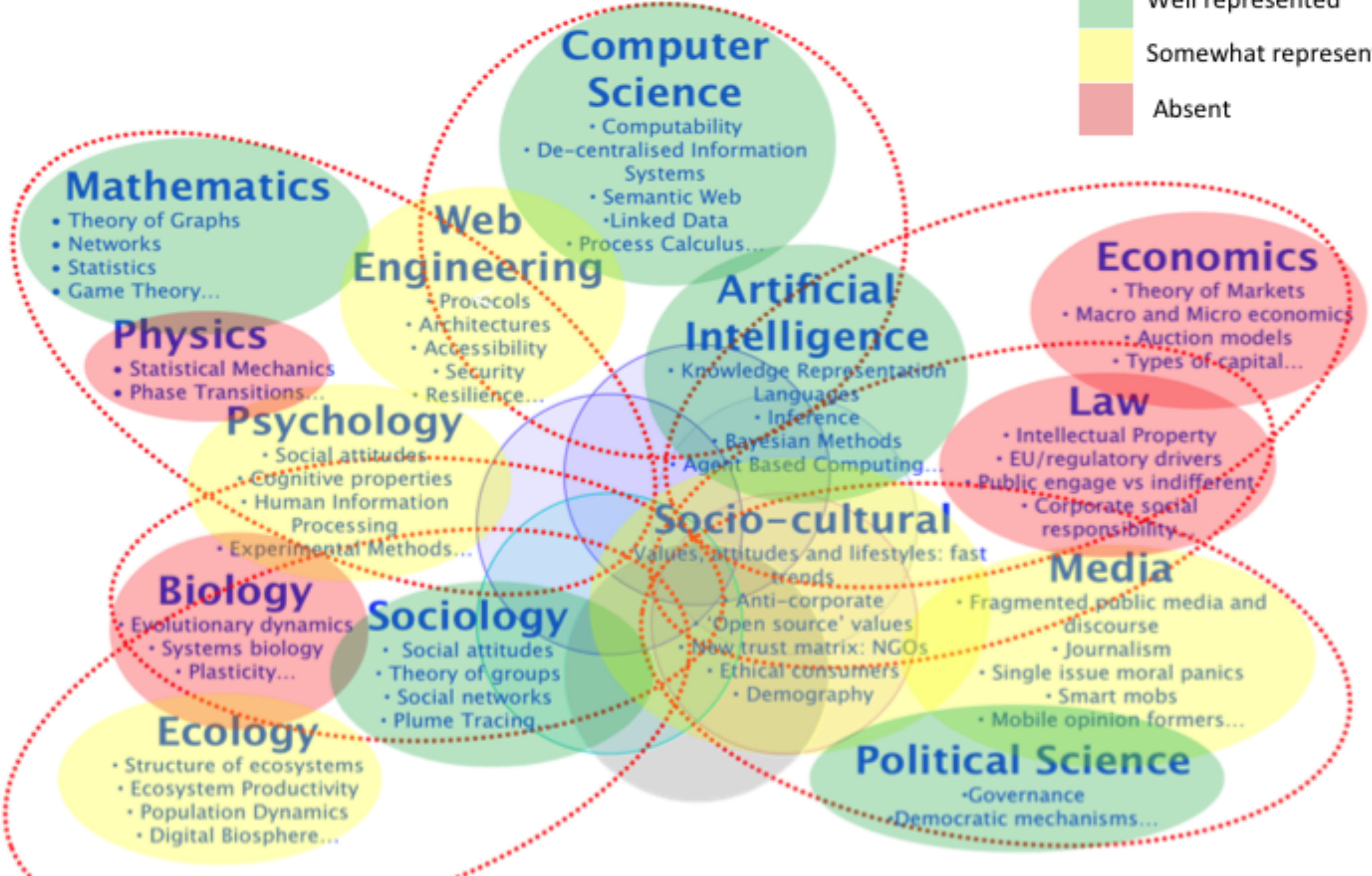
Les Carr, et.al. http://slidesha.re/142MFrV

Web Science: Components

🟩 Well represented

🟨 Somewhat represented

🟥 Absent



Computer Science
- Computability
- De-centralised Information Systems
- Semantic Web
- Linked Data
- Process Calculus...

Mathematics
- Theory of Graphs
- Networks
- Statistics
- Game Theory...

Web Engineering
- Protocols
- Architectures
- Accessibility
- Security
- Resilience...

Physics
- Statistical Mechanics
- Phase Transitions...

Artificial Intelligence
- Knowledge Representation Languages
- Inference
- Bayesian Methods
- Agent Based Computing...

Economics
- Theory of Markets
- Macro and Micro economics
- Auction models
- Types of capital...

Psychology
- Social attitudes
- Cognitive properties
- Human Information Processing
- Experimental Methods...

Law
- Intellectual Property
- EU/regulatory drivers
- Public engage vs indifferent
- Corporate social responsibility...

Socio-cultural
- Values, attitudes and lifestyles: fast trends
- Anti-corporate
- 'Open source' values
- New trust matrix: NGOs
- Ethical consumers
- Demography

Biology
- Evolutionary dynamics
- Systems biology
- Plasticity...

Sociology
- Social attitudes
- Theory of groups
- Social networks
- Plume Tracing...

Media
- Fragmented public media and discourse
- Journalism
- Single issue moral panics
- Smart mobs
- Mobile opinion formers...

Ecology
- Structure of ecosystems
- Ecosystem Productivity
- Population Dynamics
- Digital Biosphere...

Political Science
- Governance
- Democratic mechanisms...

Adapted from Shadbolt by Hooper, Marie, Kalampokis

Clare Hooper, et.al. http://bit.ly/R813sC

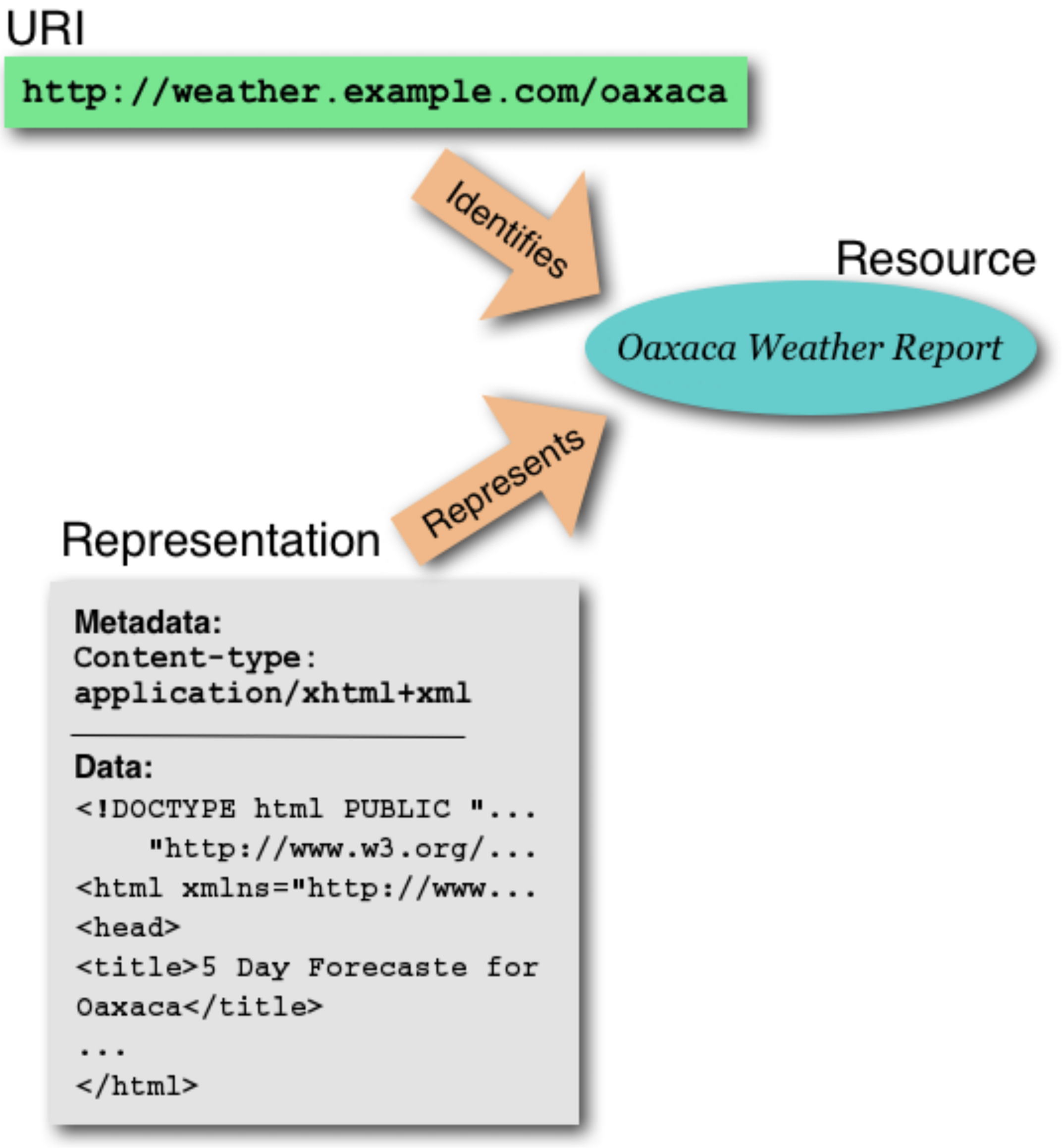
Web Architecture

It's quite simple, really! ;)

- A standard system for **identifying** resources
- Standard formats for **representing** resources
- A standard protocol for **exchanging** resources

Relevant core standards:

- URIs (URLs): Universal Resource Identifiers
- HTML: Hypertext Markup Language
- HTTP: Hypertext Transfer Protocol



URI

`http://weather.example.com/oaxaca`

Identifies

Resource

Oaxaca Weather Report

Represents

Representation

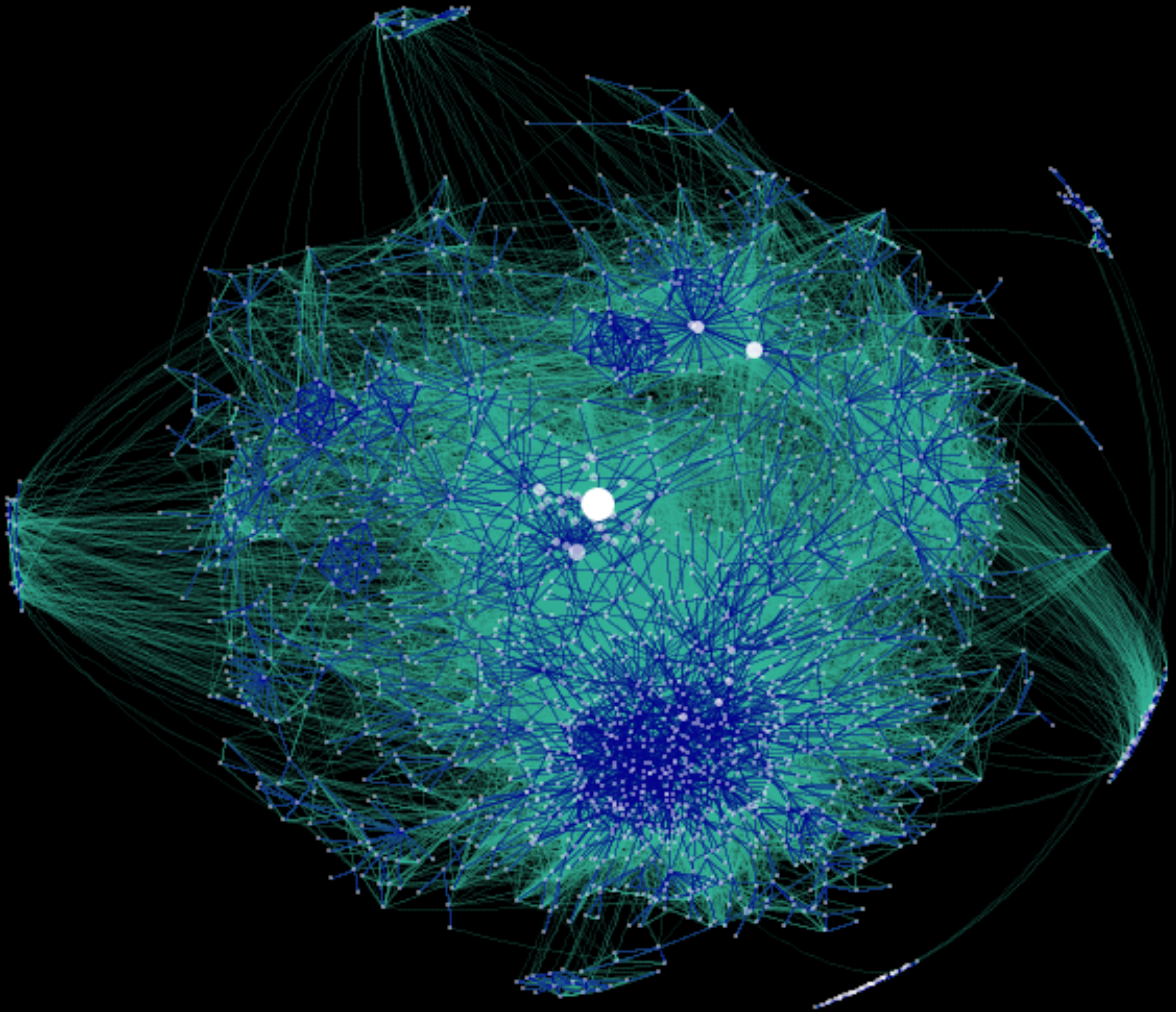
Metadata:

```
Content-type:
application/xhtml+xml
```

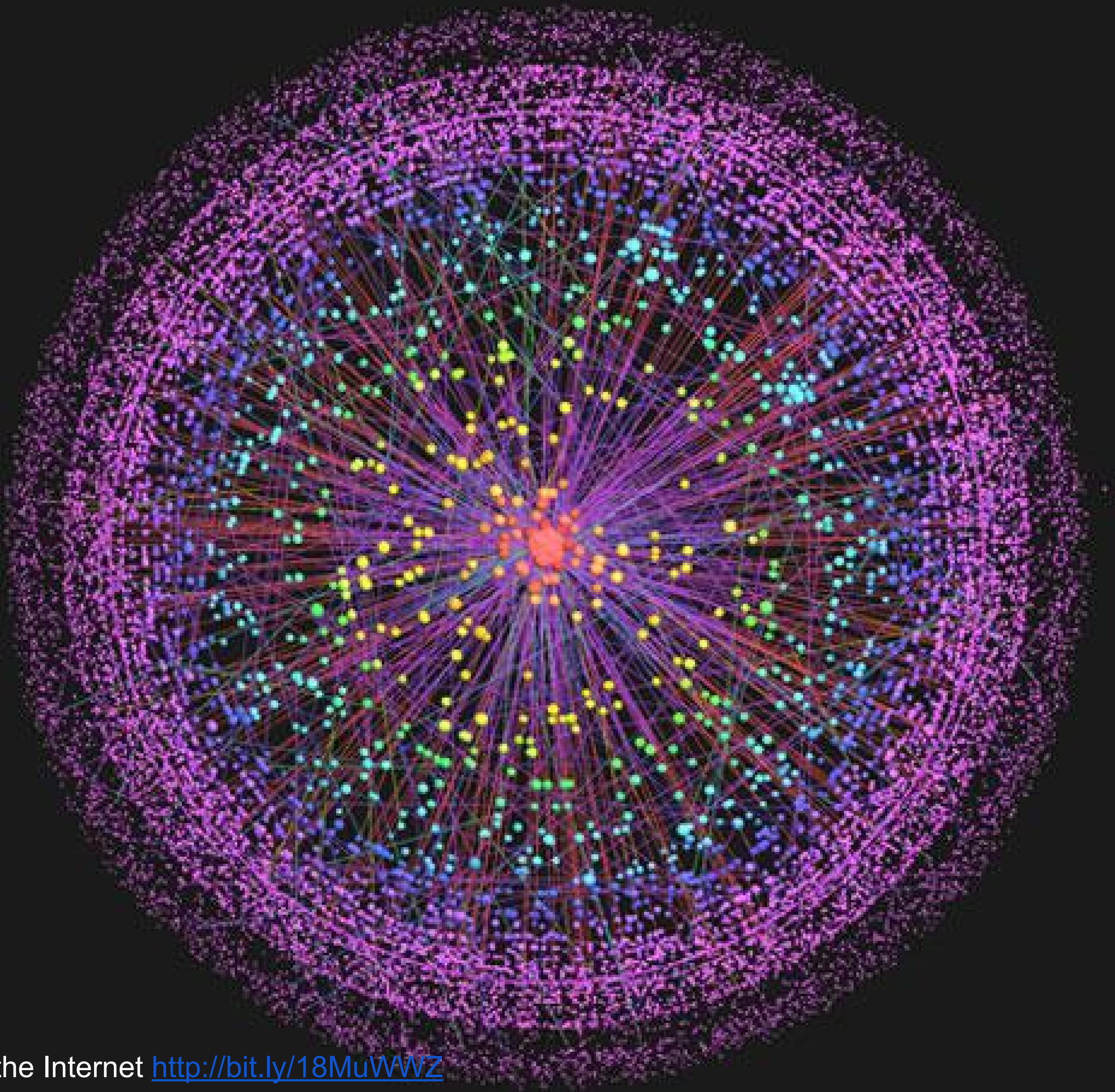
Data:

```
<!DOCTYPE html PUBLIC "...
    "http://www.w3.org/...
<html xmlns="http://www...
<head>
<title>5 Day Forecaste for
Oaxaca</title>
...
</html>
```

Architecture of the World Wide Web, Volume One http://www.w3.org/TR/webarch/



Data Mining: Mapping the Blogosphere http://bit.ly/18MuXdD



Mapping the Internet http://bit.ly/18MuWWZ

Identifying Resources (1)

- A global identification system is essential
  - to share information about resources
  - to reason about resources
  - to modify or exchange resources
- "Resources" are anything that can be linked to or spoken of
  - Documents, cat videos, people, ideas...
- Not all resources are "on" the Web
  - They might be referenced from the Web...
  - ...while not being retrievable from it
  - These are (so called) "information resources"

Les Carr, et.al. http://slidesha.re/142MFrV

Identifying Resources (2)

- A global standard is required; the **URI** is it
- Others systems are possible...
  - ...but added value of a **single global system** of identifiers is high
  - Enables linking, bookmarking and other functions across heterogeneous applications
- How are URI used?
  - All resources have URIs associated with them
  - Each URI identifies a single resource in a context-independent manner
  - URIs act as names and (usually) addresses
  - In general URIs are "opaque"

Uniform Resource Identifier (URI): Generic Syntax (RFC 3986) http://www.ietf.org/rfc/rfc3986.txt

Identifying Resources (4)

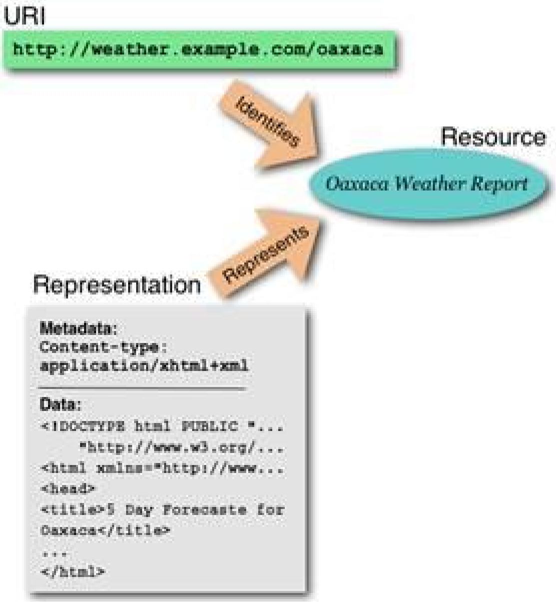
- "URIs *identify* and URLs *locate...*"
  - ...and identify
- URLs are URIs aligned with protocols
  - URLs include the "access mechanism" or "network location", e.g. http:// or ftp://
  - How to "dereference" the URI and retrieve the thing
- URL examples
  - ftp://ftp.is.co.za/rfc/rfc1808.txt
  - http://www.ietf.org/rfc/rfc2396.txt
  - mailto:John.Doe@example.com
  - telnet://192.0.2.16:80/

Uniform Resource Identifier (URI): Generic Syntax (RFC 3986) http://www.ietf.org/rfc/rfc3986.txt

Representing Resources (1)

- Resources are manifest as digital files
- The Web recognizes a (growing) set of file formats
  - The original and workhorse is HTML...
  - ...but there are many others
- Retrievable resources on the web serve multiple purposes
  - Resources **encode** information and data
  - Resources **aggregate links** to other resources
- This is what makes The Web(tm) a "web..."

Resources (nodes) aggregate links to other resources to create a Web



URI

http://weather.example.com/oaxaca

Identifies

Resource

Oaxaca Weather Report

Represents

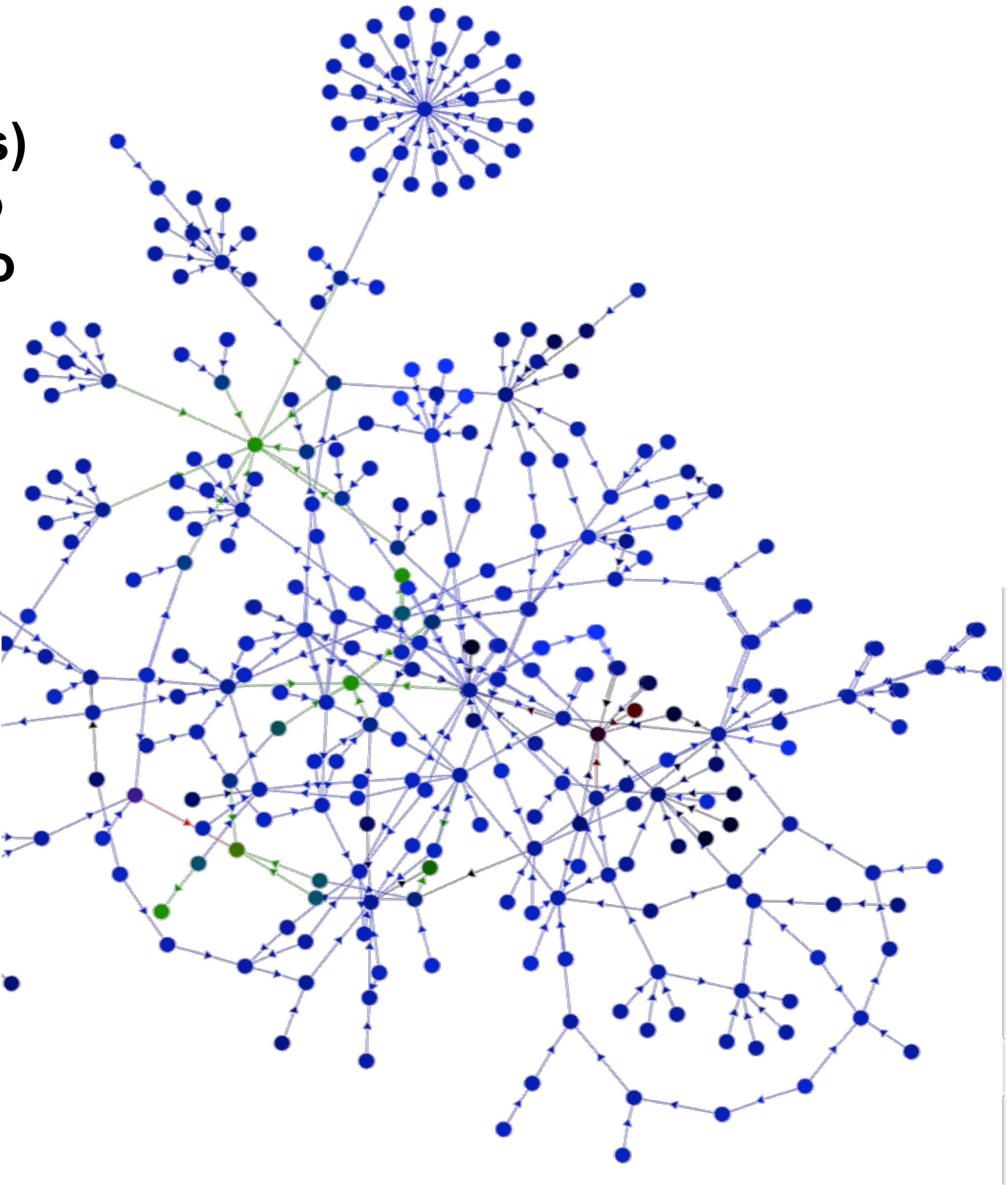
Representation

Metadata:

```
Content-type:
application/xhtml+xml
```

Data:

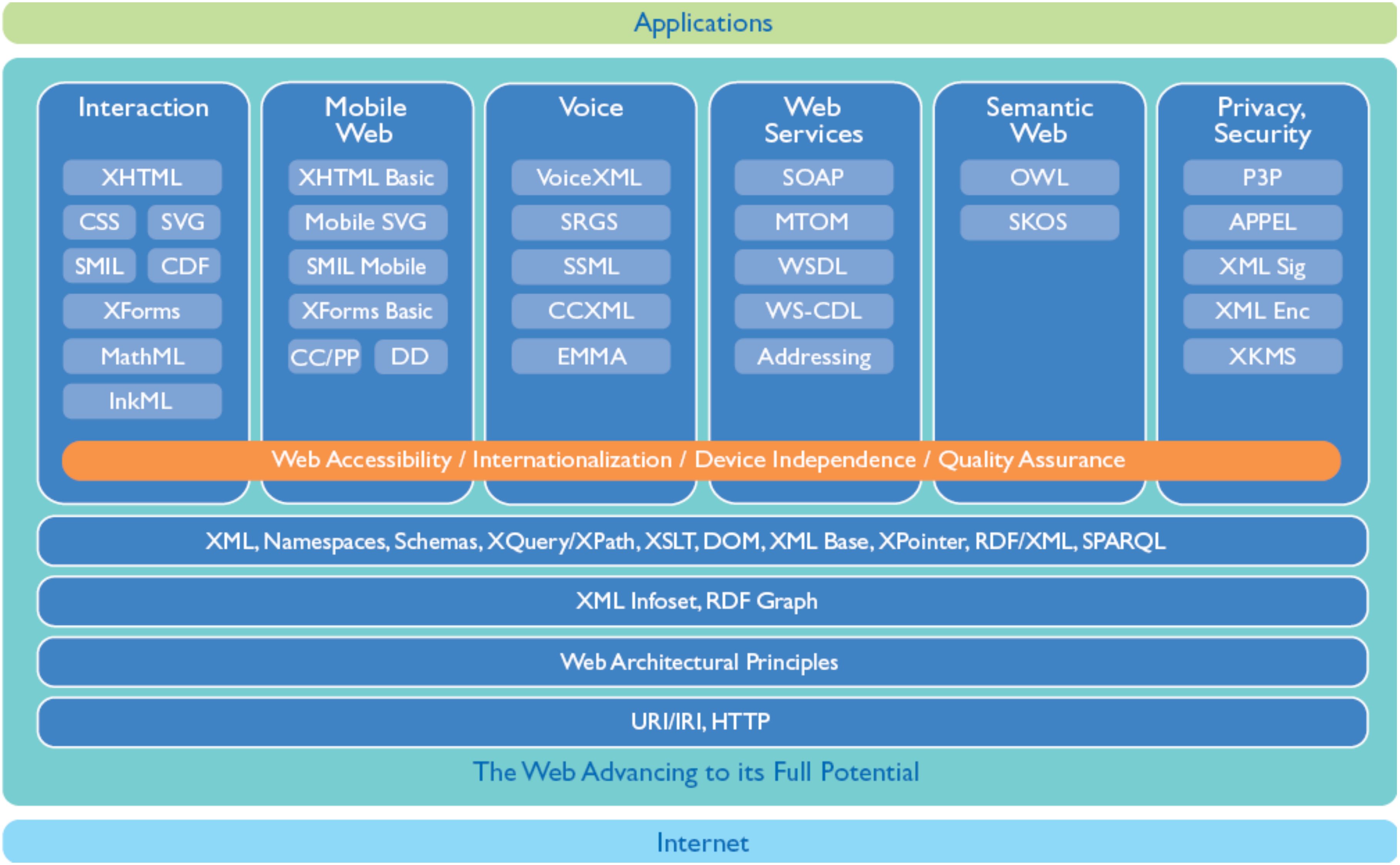
```
<!DOCTYPE html PUBLIC "...
    "http://www.w3.org/...
<html xmlns="http://www...
<head>
<title>5 Day Forecaste for
Oaxaca</title>
...
</html>
```



Retrieving Resources (1)

- **Review:** URIs that reference retrievable resources -- URLs -- must specify a protocol for retrieval
- The original and most common Web protocol is HTTP
- Specialized protocols are possible but resources may appear "off the grid..."

URIs, HTTP, many formats...



Principles for creating a healthy Web

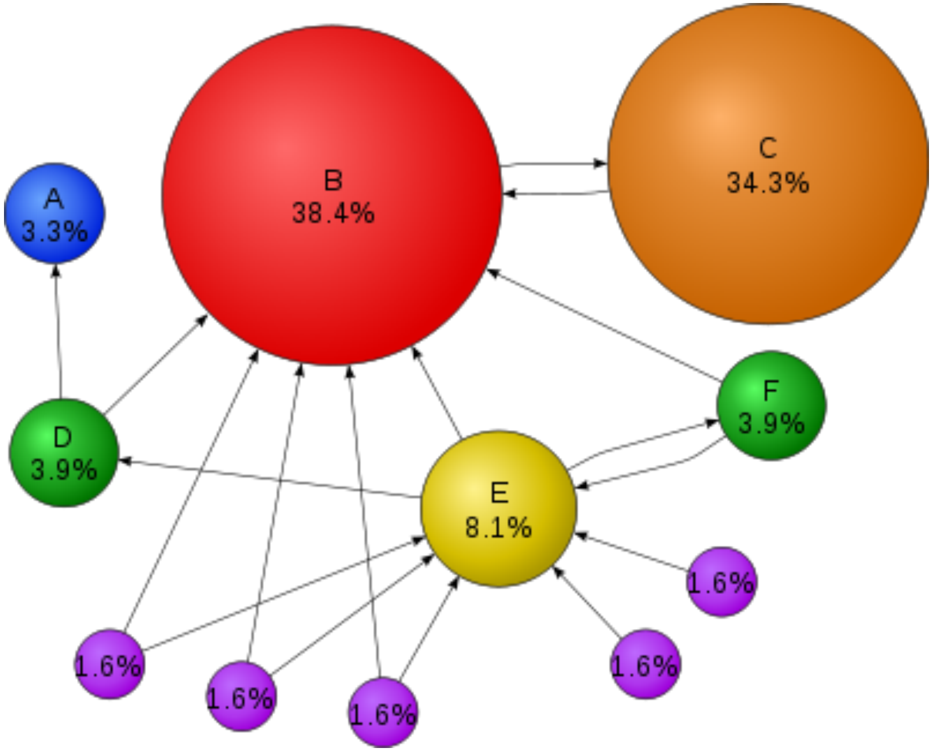
- Use URIs as names for things
- Use HTTP URIs so people can "look up" those names
- When someone "looks up" a URI, return useful information
  - use the standards to do it
- Include links to other URIs, so the Consumer can discover more things
  - People or applications

Why is linking important???

Tim Berners-Lee http://www.w3.org/DesignIssues/LinkedData.html

Implications of a well-connected Web: Google PageRank

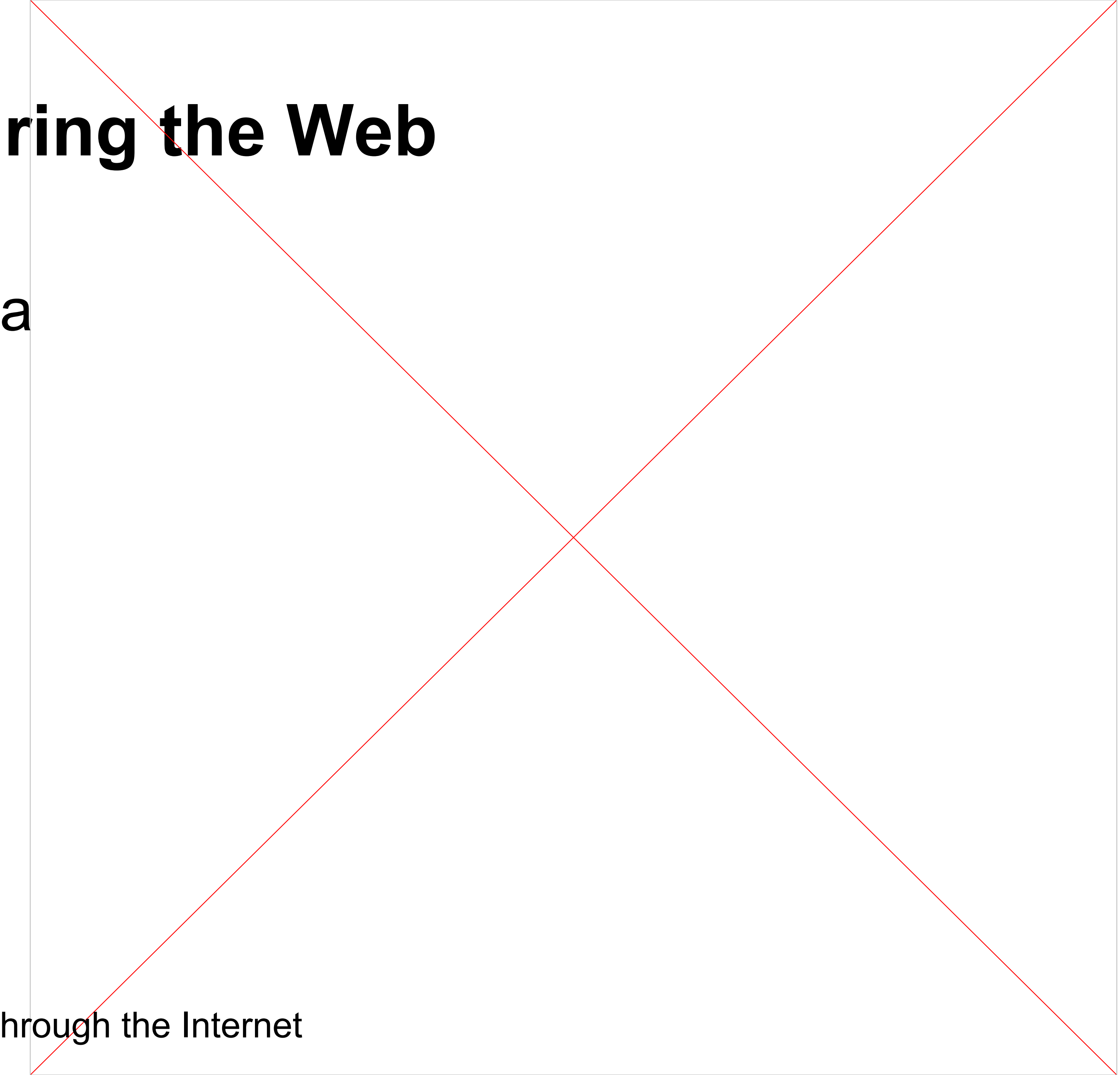
- Links to other nodes as a "vote" of quality and/or relevance



PageRank https://en.wikipedia.org/wiki/PageRank

Measuring the Web

Web as a Network



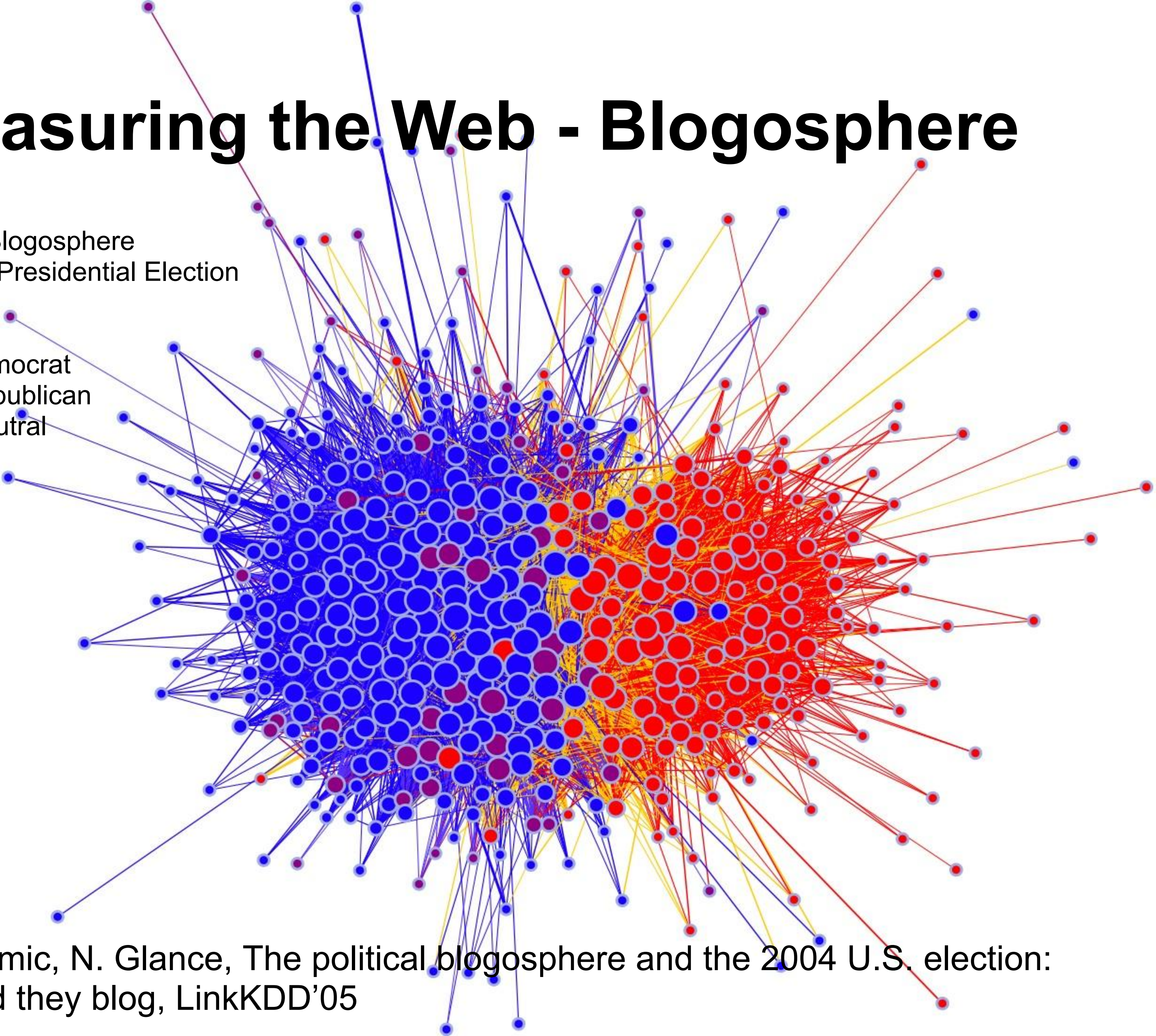
Router network through the Internet

Measuring the Web

- The rich variety of networks on the Web
  - Router network
  - Web page network (linking via hyperlinks)
  - Document network (citation network on DBLP*, etc.)
  - Social networks
    - Facebook: friendship, comment-reply, tag, and all kinds of social relationship on Facebook
    - Twitter: follower, retweet, mention, reply, etc.
    - Blogosphere: friendship, visiting, comment, etc.
    - LinkedIn: colleague, classmate, etc.
    - Crowdsourcing: collaboration, co-worker, etc.
    - Other social media...

*The DBLP Computer Science Bibliography http://www.informatik.uni-trier.de/~ley/db/

Measuring the Web - Blogosphere

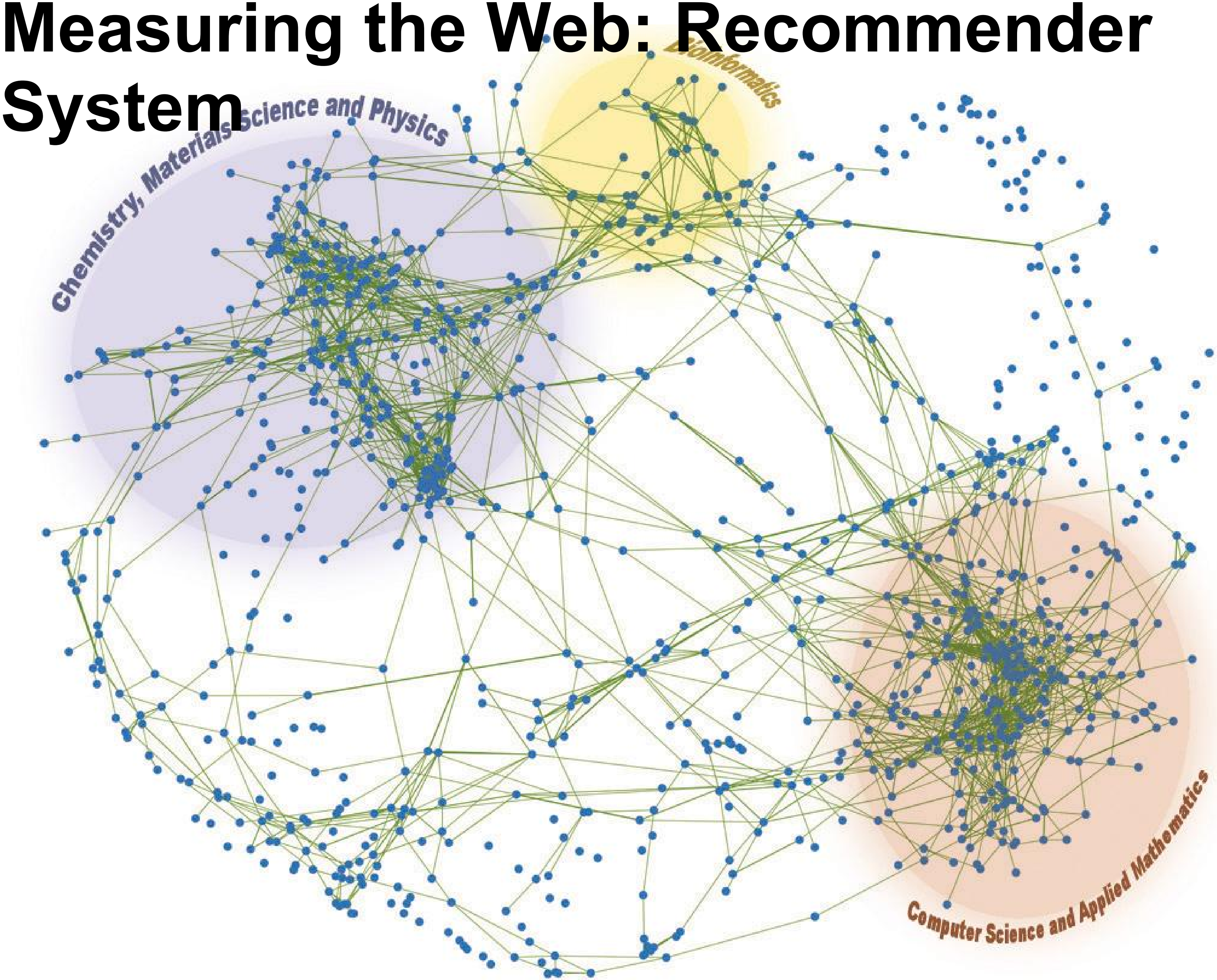


Political Blogosphere
2004 US Presidential Election

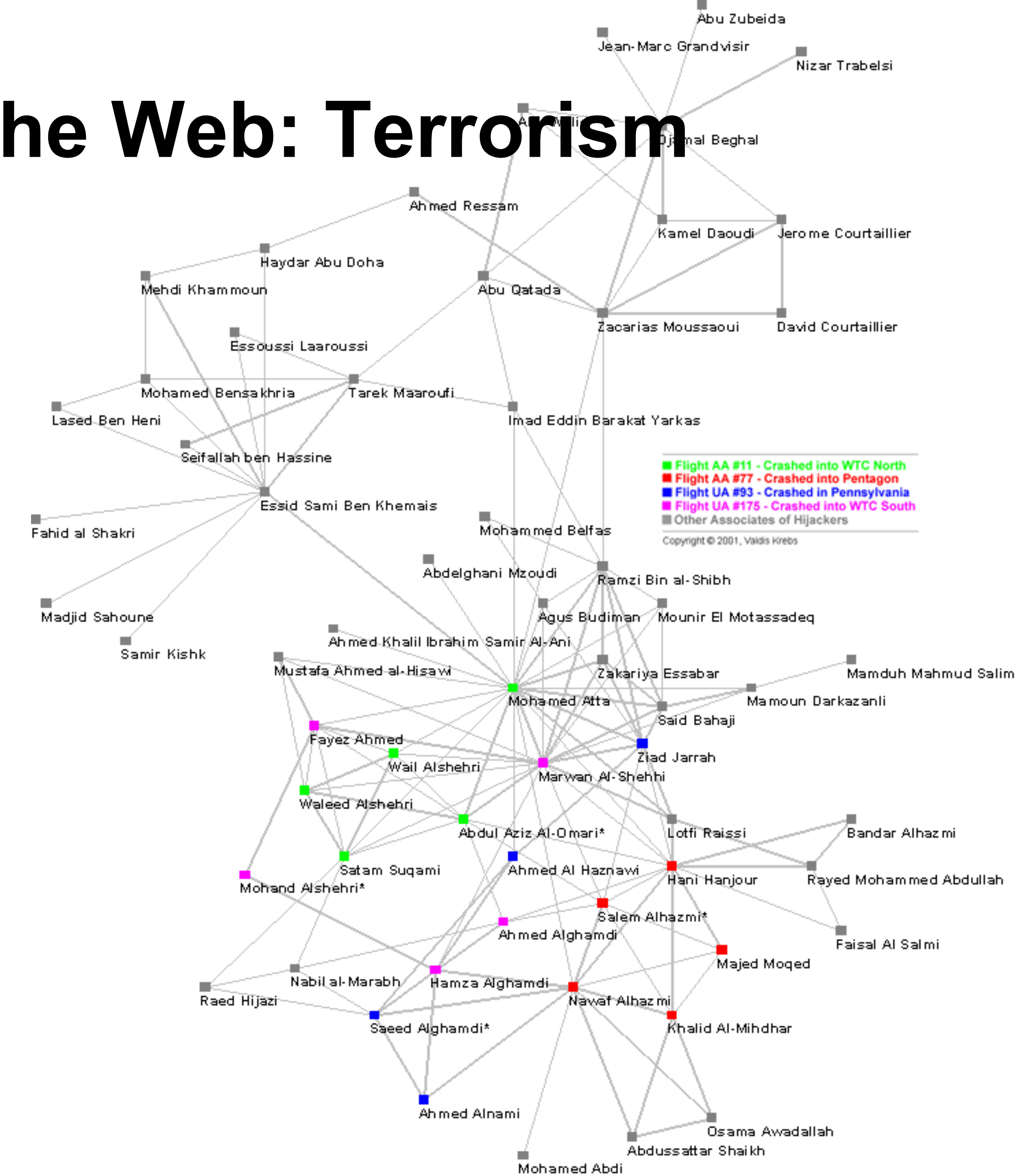
Bloggers:
Blue -Democrat
Red - Republican
Pink - Neutral

L. Adamic, N. Glance, The political blogosphere and the 2004 U.S. election: divided they blog, LinkKDD'05

Measuring the Web: Recommender System


Chemistry, Materials Science and Physics
Bioinformatics
Computer Science and Applied Mathematics

Measuring the Web: Terrorism

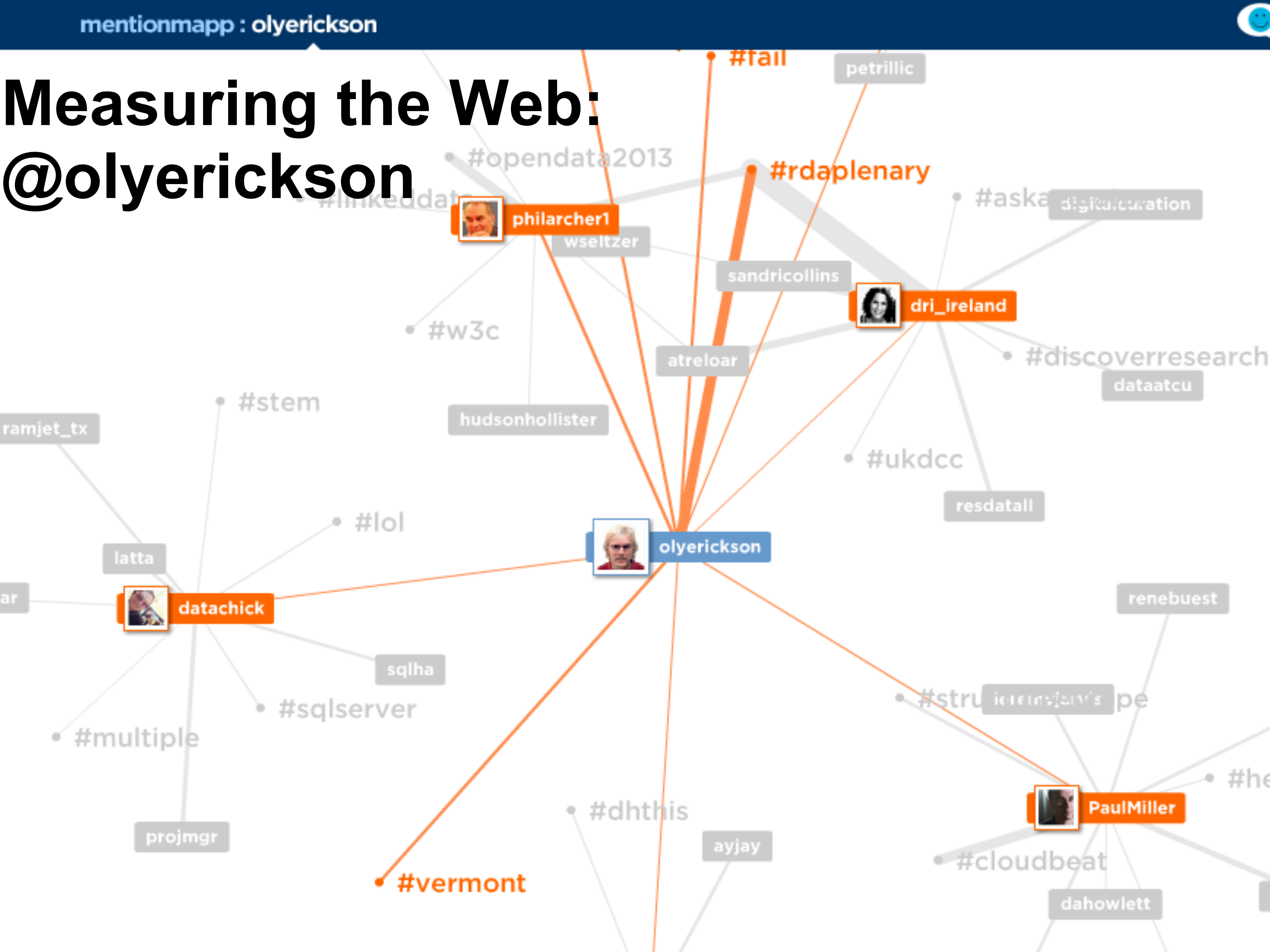


Copyright © 2001, Valdis Krebs

Measuring the Web: Twitter

M. D. Conover, et al. Political Polarization on Twitter, ICWSM’11

Measuring the Web: @olyerickson



Analyzing networks on the Web

Measure...

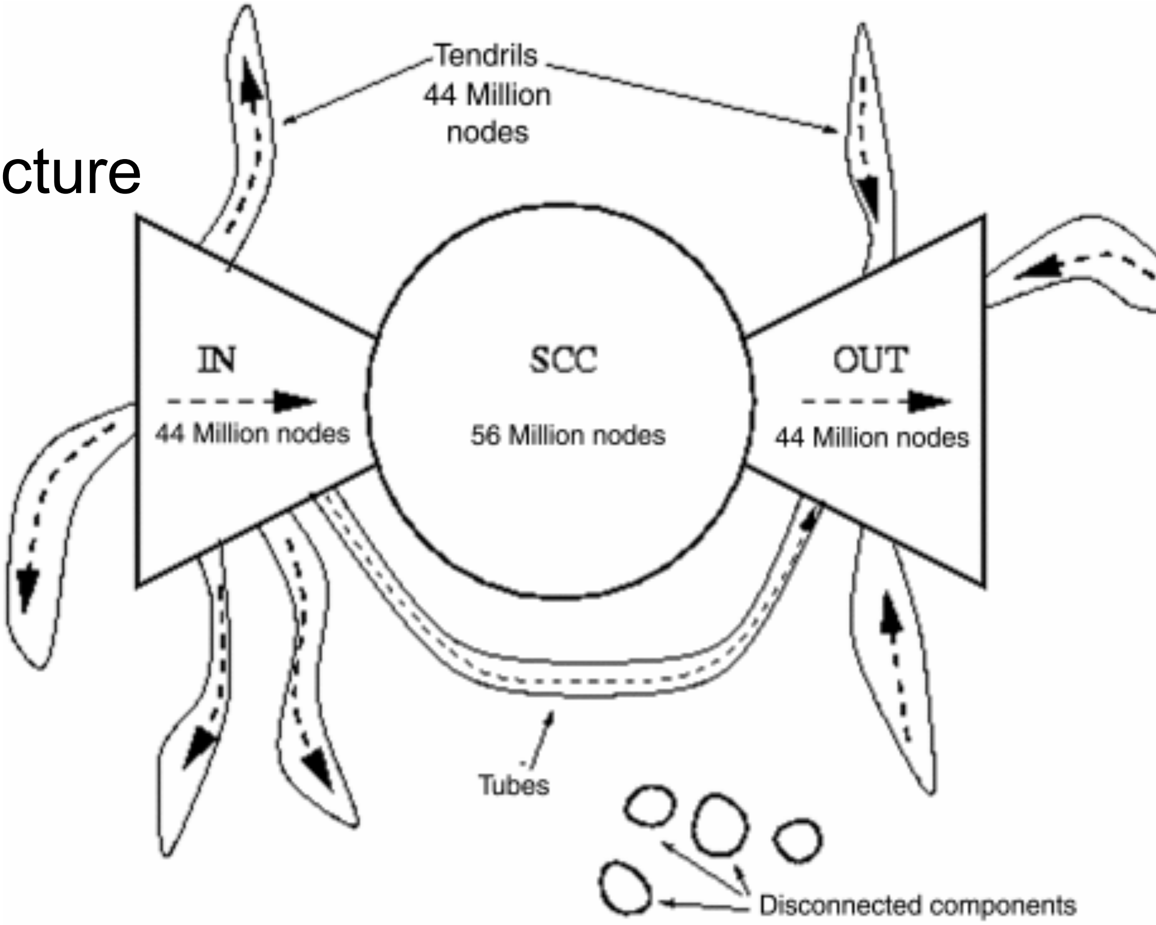
- # of nodes
- # of edges
- Diameter and radius
- Network density
- Degree distribution
- Clustering coefficient
- Average shortest path length
- Strongly/weakly connected components
- Betweenness/Closeness centrality
- **Bow-tie structure**
- Community discovery
- Key nodes discovery
- etc...

Measuring the Web

“Bow-tie” structure

Overall view of the structure of the Web

- SCC
- IN
- OUT
- Tendrils
- Tubes
- Disconnected



A Broder et al. Graph structure in the Web, Computer Networks, (2000)

Measuring the Web

- It's a “Small World” after all...
  - Most pairs of pages separated by small # of links
  - Almost always by fewer than 20 links
  - "Diameter" of central core is 28, very small compared to the size of the Web
  - Analysis suggests diameter will grow logarithmically with the size of the Web (ie slowly)
  - Diameter of social networks decreases over time
- Conclusion: The Web is “smaller” than we thought!
- “Six degrees of separation” verified in Social Web

R. Albert, H. Jeong and A.-L. Barabasi, Diameter of the World Wide Web, Nature 401 (1999) 130–131. http://bit.ly/18atsYA

J. Leskovec, etc. Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations, KDD (2005)

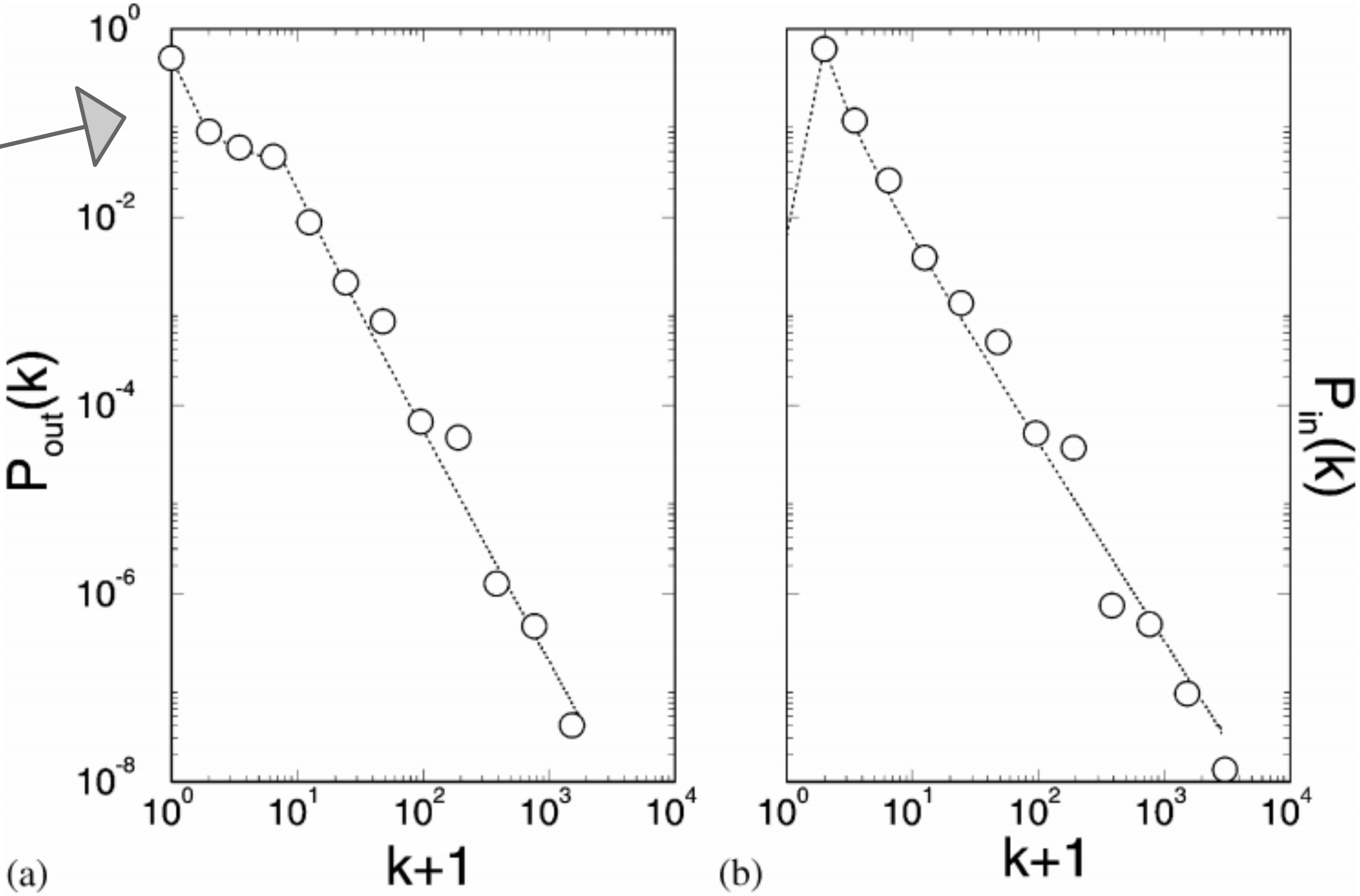
Measuring the Web

"Scale-free" property

$$P(k) \sim k^{-\gamma}$$

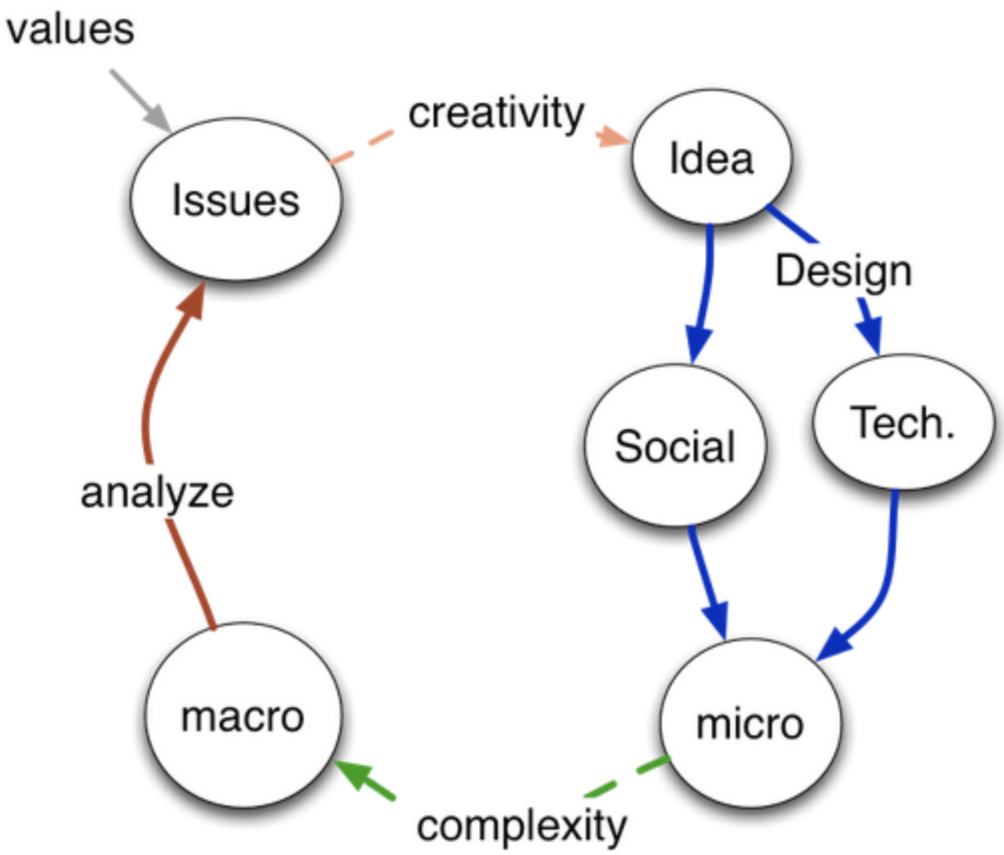
Highly Connected Hubs

"Rich get richer"



A live model: http://ccl.northwestern.edu/netlogo/models/PreferentialAttachment

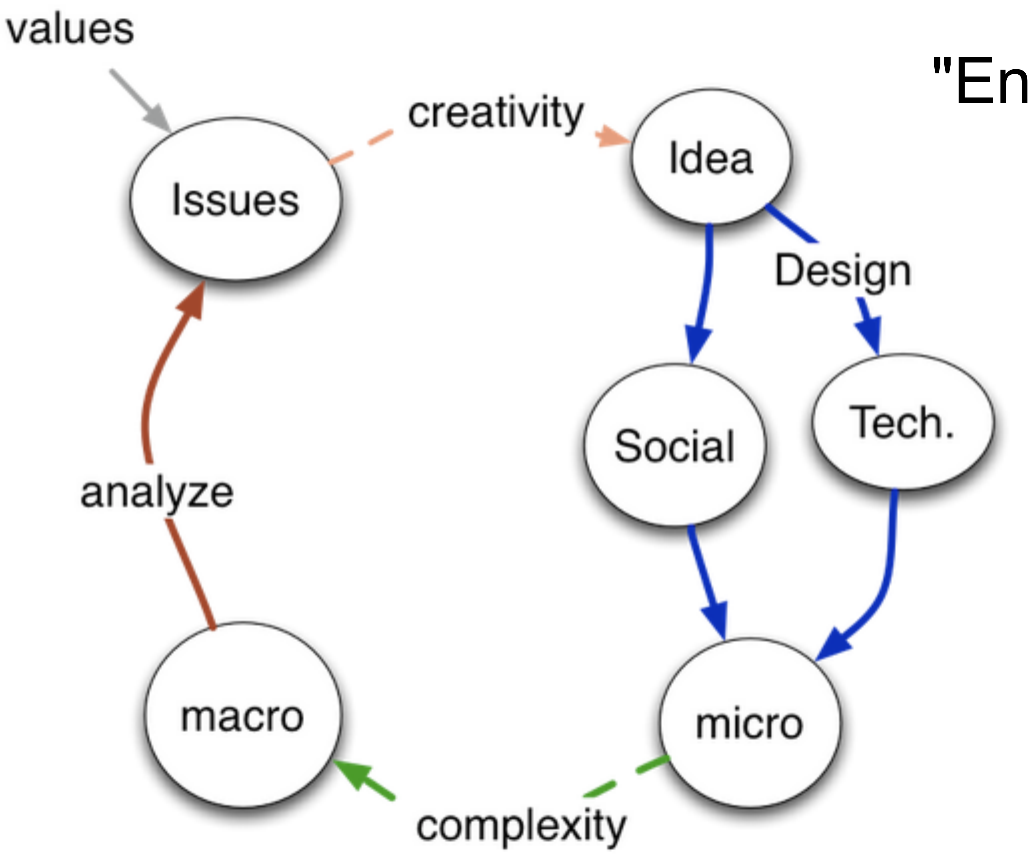
The Web Science Method



Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(10)

The Web Science Method

"Engineering"



values

Issues

creativity

Idea

Design

Social

Tech.

micro

complexity

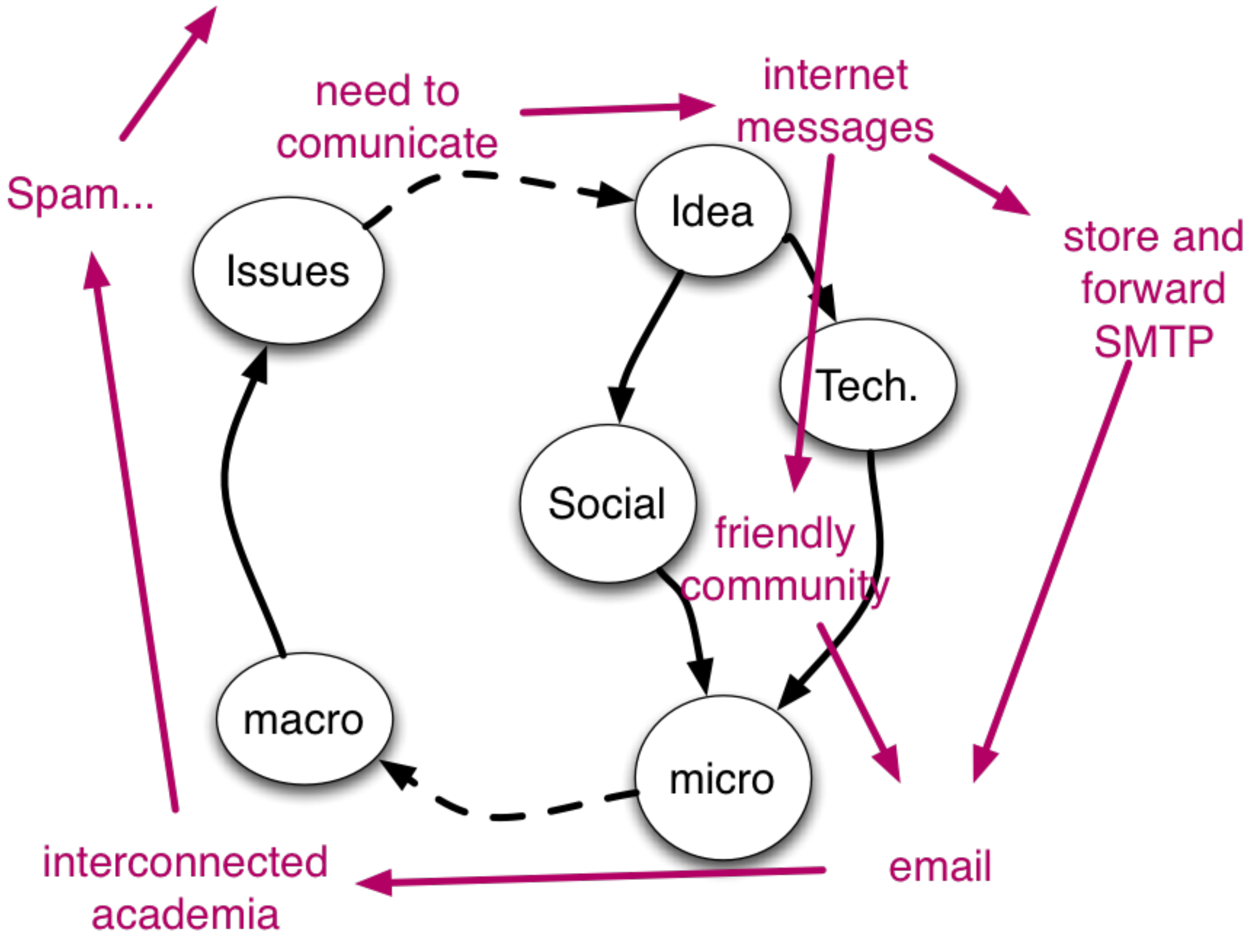
macro

analyze

"Science"

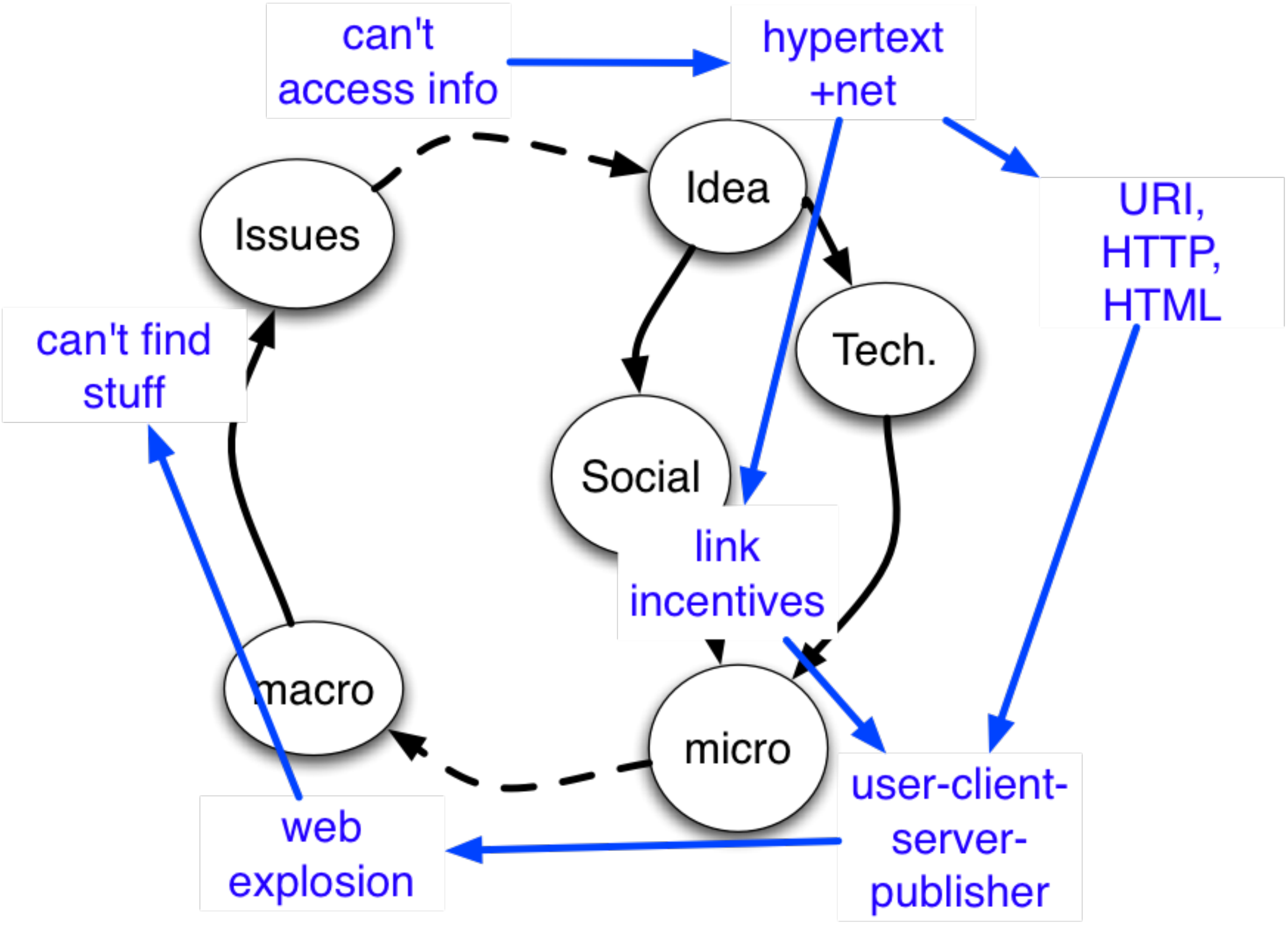
Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(10)

Applied to email...



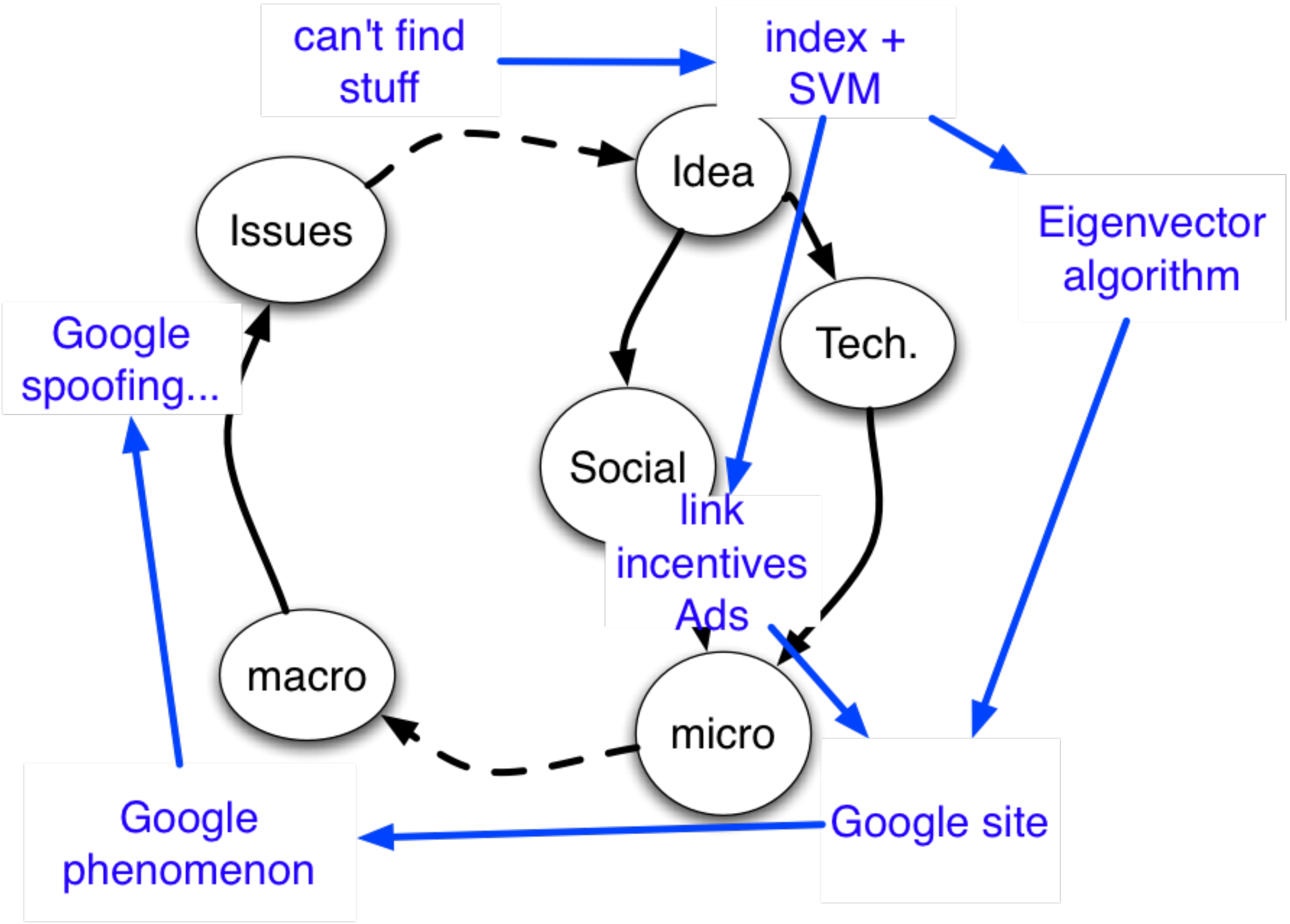
Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(16)

Applied to the original Web...



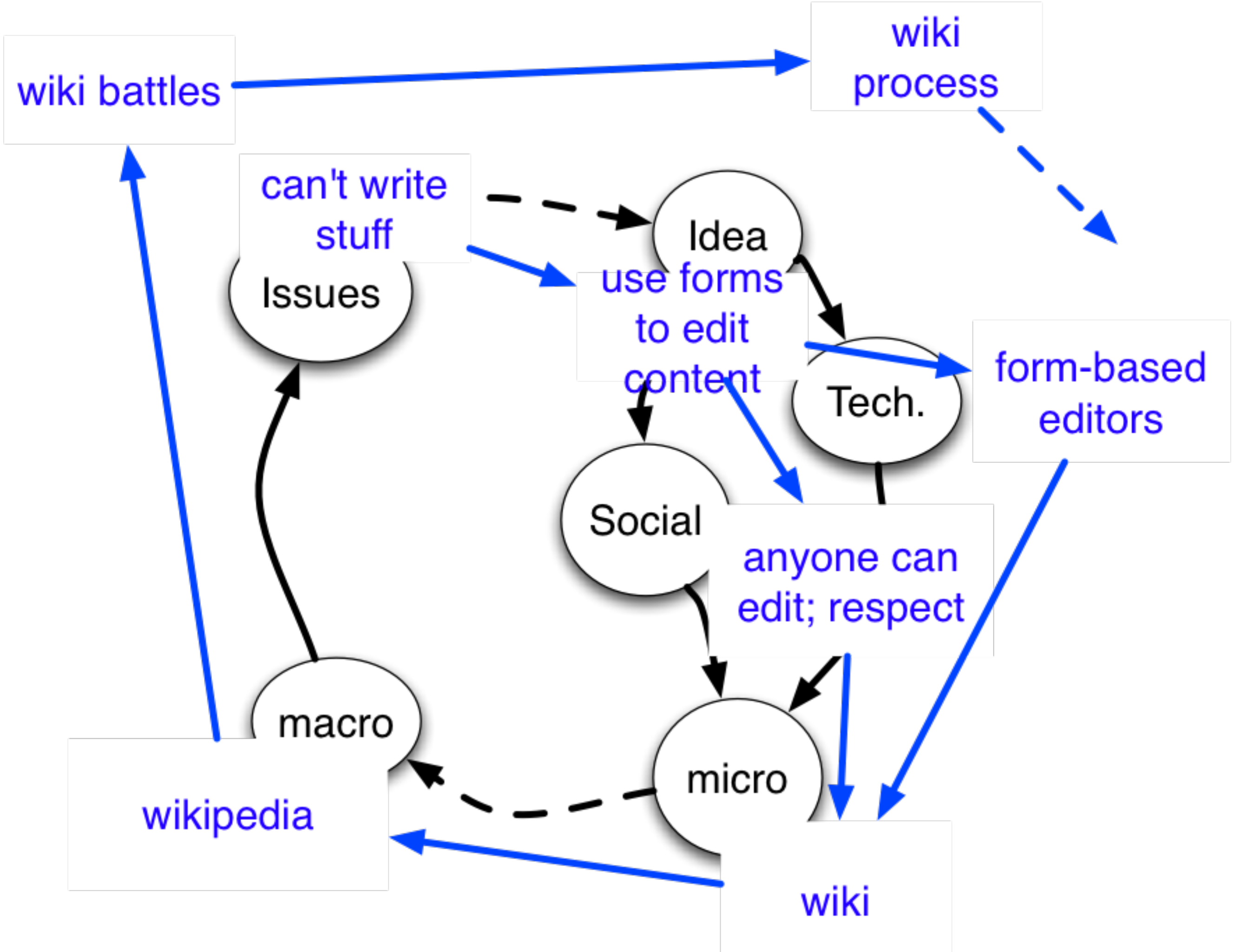
Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(18)

Applied to Google's Web...



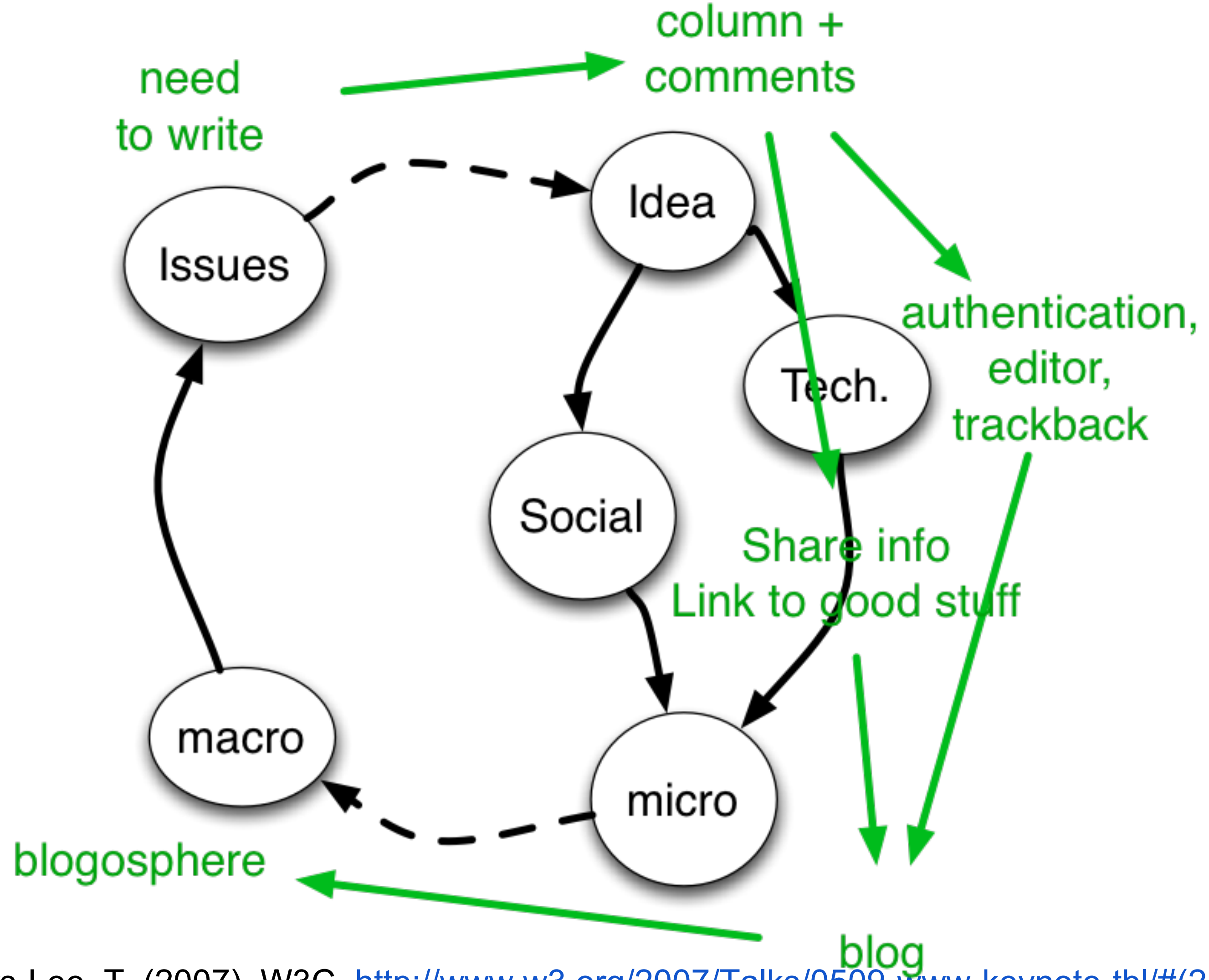
Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(19)

Applied to Wikis...



Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(20)

Applied to Blogs...



Berners-Lee, T. (2007). W3C. http://www.w3.org/2007/Talks/0509-www-keynote-tbl/#(21)

Social Aspects of the Web

"Visual complexity produces opacity. Massive individualizing data produces beautiful, playful hairballs which show us nothing."

- Bruno Latour @ CHI2013

For discussion see, "What baboon notebooks, monads, state surveillance and network diagrams have in common: Bruno Latour at CHI2013" http://bit.ly/14Y3d3u

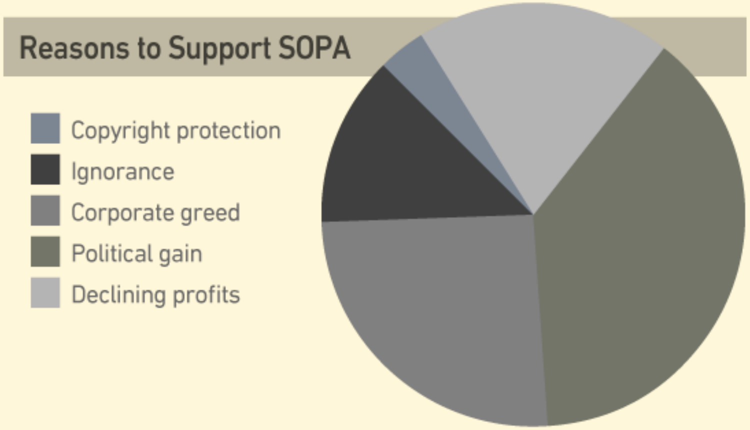
Multiple disciplines, multiple methods

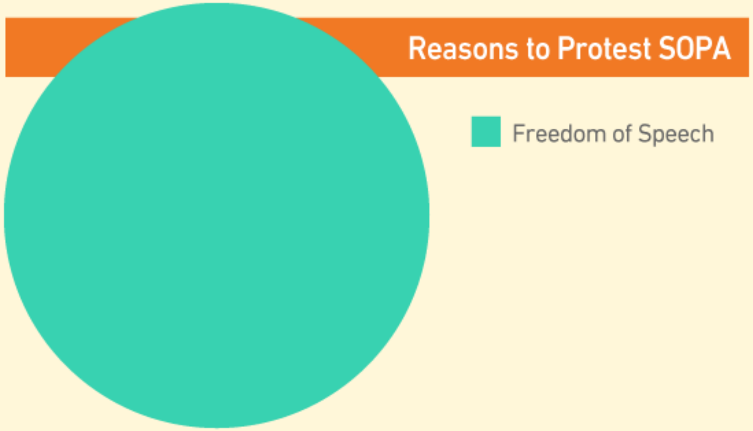
- Given it’s multiple disciplines, the argument is for a mixed-methods approach to measuring the web. This means both **quantitative AND qualitative** methods should be employed by researchers.
    - Pros: More robust, comprehensive understanding of “human social behavior”
    - Cons: Diametrically opposed philosophies in data gathering and analysis
- Unanswered questions:
    - Replicability
    - Bias
    - Objectivity and Accuracy
- Ethics

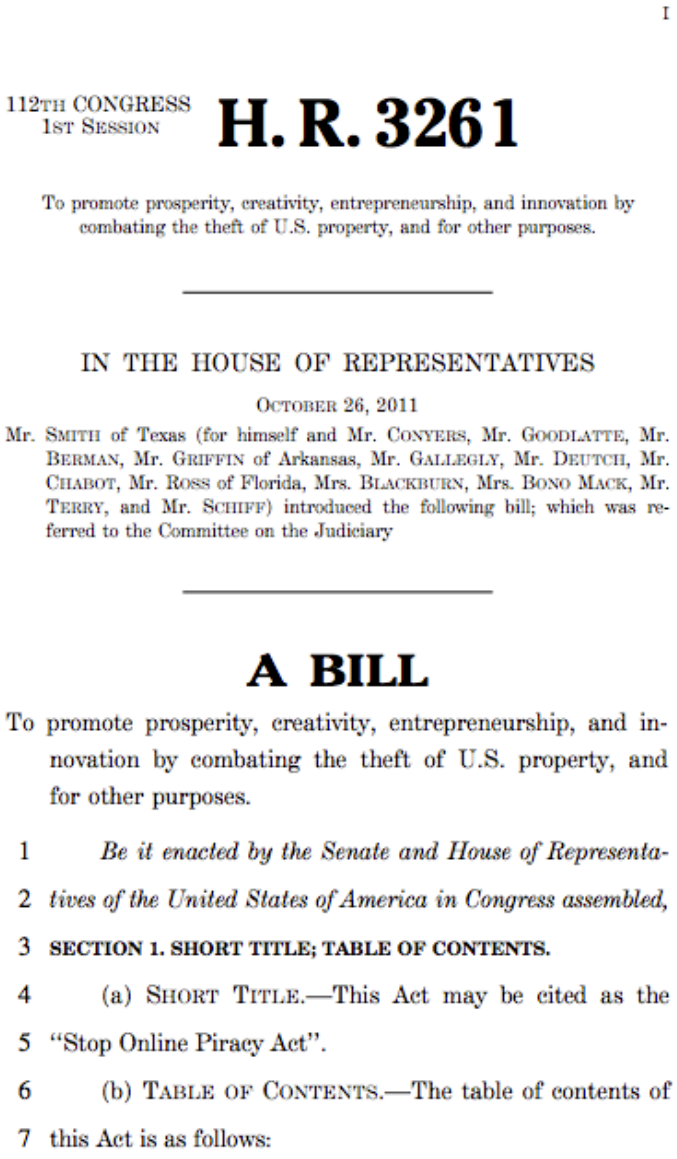
Web Governance, Security and Standards

Who should govern the Web?

SOPA Stop Online Piracy Act







112TH CONGRESS
1ST SESSION

H. R. 3261

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

OCTOBER 26, 2011

Mr. SMITH of Texas (for himself and Mr. CONYERS, Mr. GOODLATTE, Mr. BERMAN, Mr. GRIFFIN of Arkansas, Mr. GALLEGLY, Mr. DEUTCH, Mr. CHABOT, Mr. ROSS of Florida, Mrs. BLACKBURN, Mrs. BONO MACK, Mr. TERRY, and Mr. SCHIFF) introduced the following bill; which was referred to the Committee on the Judiciary

A BILL

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE; TABLE OF CONTENTS.

(a) SHORT TITLE.—This Act may be cited as the "Stop Online Piracy Act".

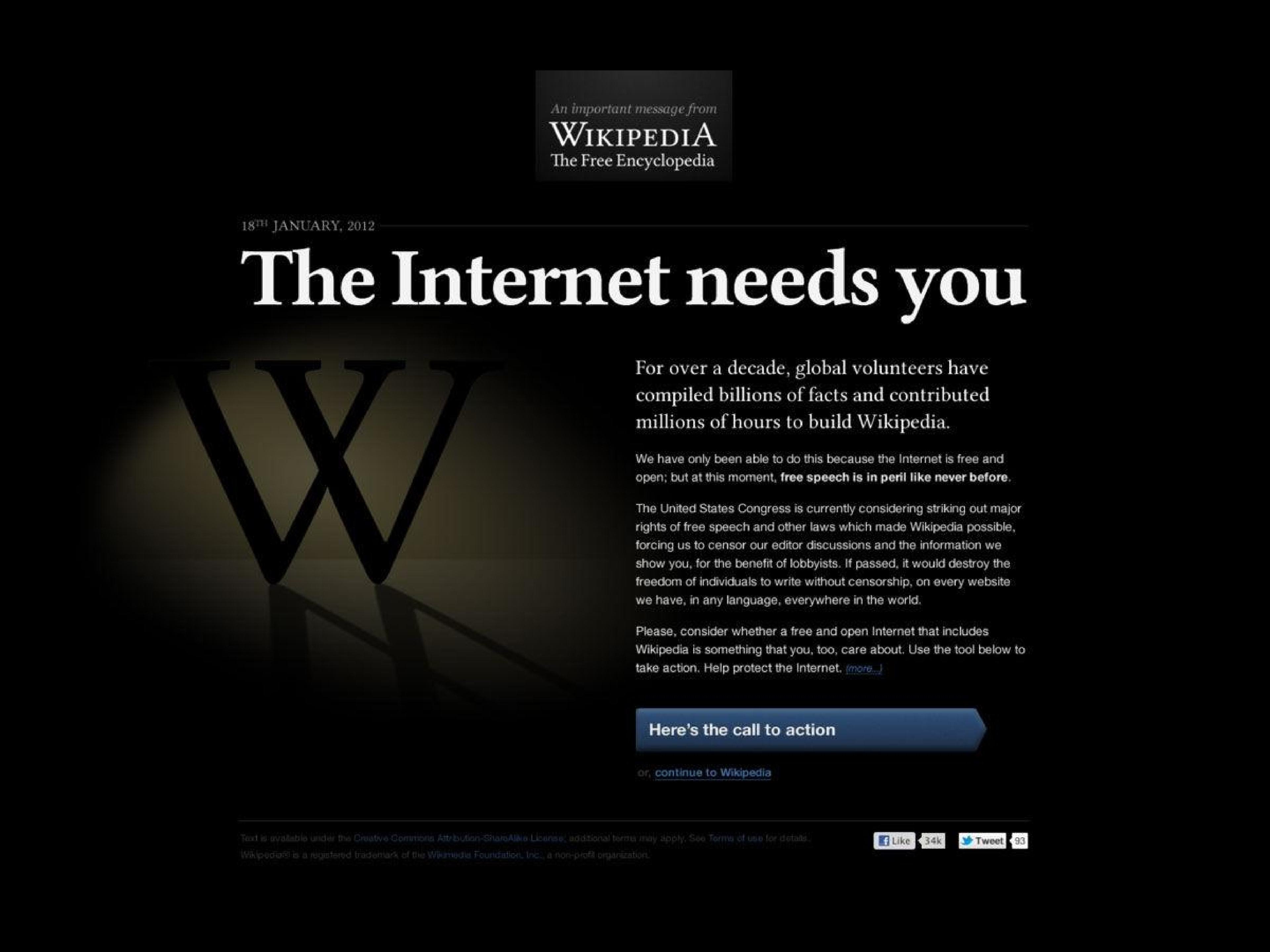
(b) TABLE OF CONTENTS.—The table of contents of this Act is as follows:

SUPPORT SOPA

News Corporation
ESPN
MARVEL
viacom
UNIVERSAL
UNIVERSAL MUSIC GROUP
MasterCard
RIAA
VISA
L'ORÉAL
NBCUniversal
MOTION PICTURE ASSOCIATION OF AMERICA

OPPOSE SOPA

Google
Aol.
YAHOO!
facebook
twitter
reddit
zynga
ebay
REPORTERS WITHOUT BORDERS
mozilla
4chan



An important message from

WIKIPEDIA

The Free Encyclopedia

18TH JANUARY, 2012

The Internet needs you

For over a decade, global volunteers have compiled billions of facts and contributed millions of hours to build Wikipedia.

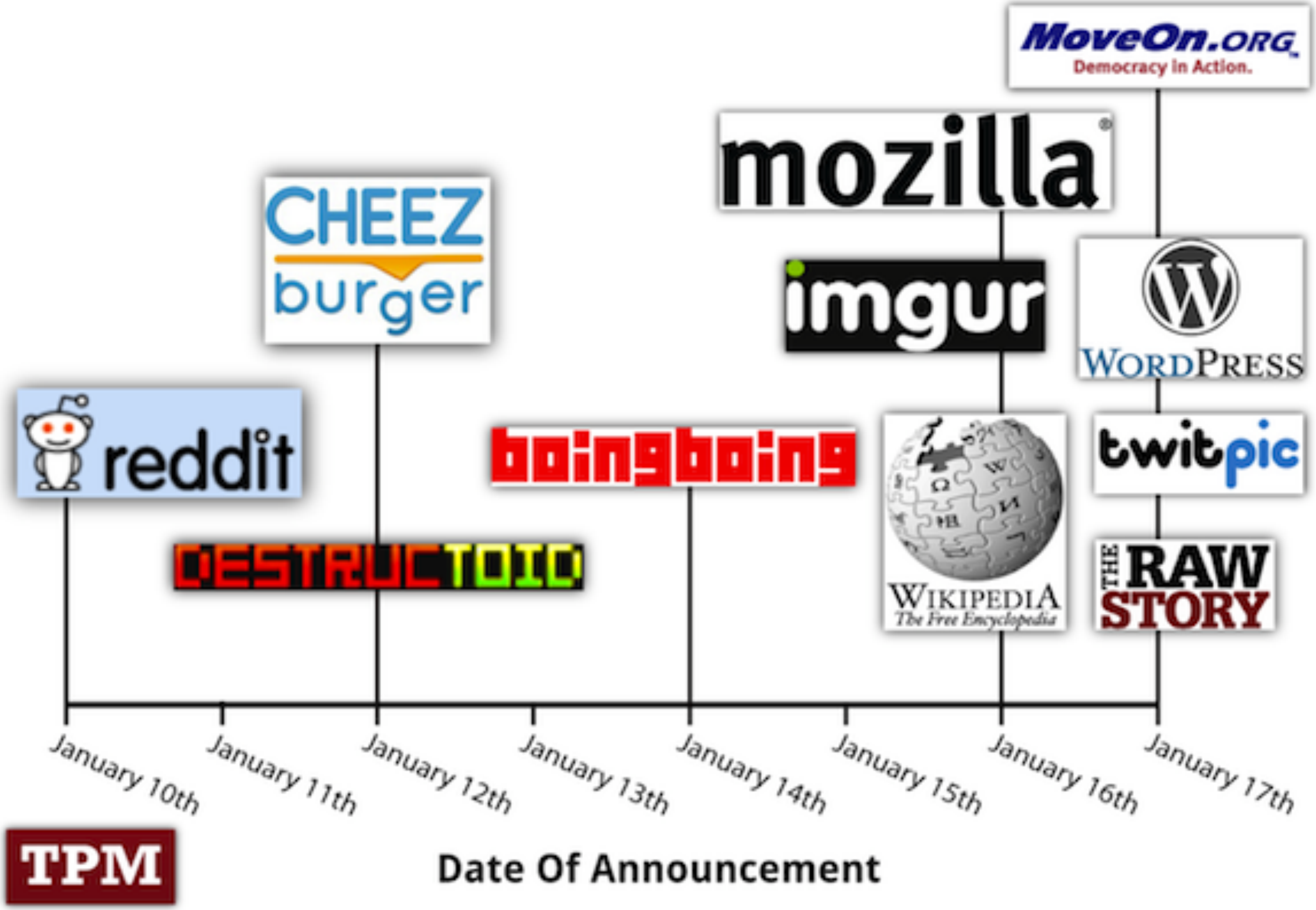
We have only been able to do this because the Internet is free and open; but at this moment, **free speech is in peril like never before**.

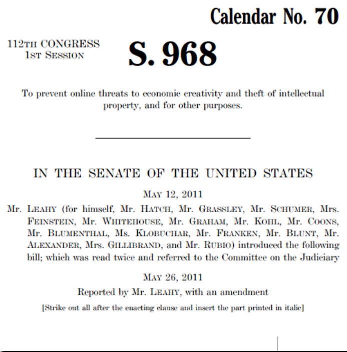
The United States Congress is currently considering striking out major rights of free speech and other laws which made Wikipedia possible, forcing us to censor our editor discussions and the information we show you, for the benefit of lobbyists. If passed, it would destroy the freedom of individuals to write without censorship, on every website we have, in any language, everywhere in the world.

Please, consider whether a free and open Internet that includes Wikipedia is something that you, too, care about. Use the tool below to take action. Help protect the Internet. *(more...)*

Websites Planning To Protest SOPA And PIPA

Timeline of Major Sites Announcing They Would 'Go Dark' On January 18





Calendar No. 70

112TH CONGRESS
1ST SESSION

S. 968

To prevent online threats to economic creativity and theft of intellectual property, and for other purposes.

IN THE SENATE OF THE UNITED STATES

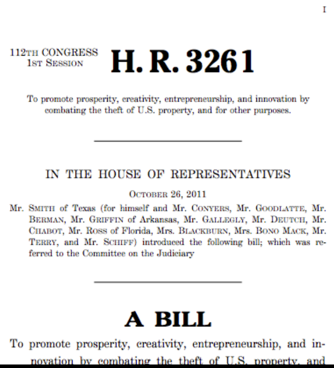
MAY 12, 2011

Mr. LEAHY (for himself, Mr. HATCH, Mr. GRASSLEY, Mr. SCHUMER, Mrs. FEINSTEIN, Mr. WHITEHOUSE, Mr. GRAHAM, Mr. KOHL, Mr. COONS, Mr. BLUMENTHAL, Ms. KLOBUCHAR, Mr. FRANKEN, Mr. BLUNT, Mr. ALEXANDER, Mrs. GILLIBRAND, and Mr. RUBIO) introduced the following bill; which was read twice and referred to the Committee on the Judiciary

MAY 26, 2011

Reported by Mr. LEAHY, with an amendment

[Strike out all after the enacting clause and insert the part printed in italic]



I

112TH CONGRESS
1ST SESSION

H. R. 3261

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

OCTOBER 26, 2011

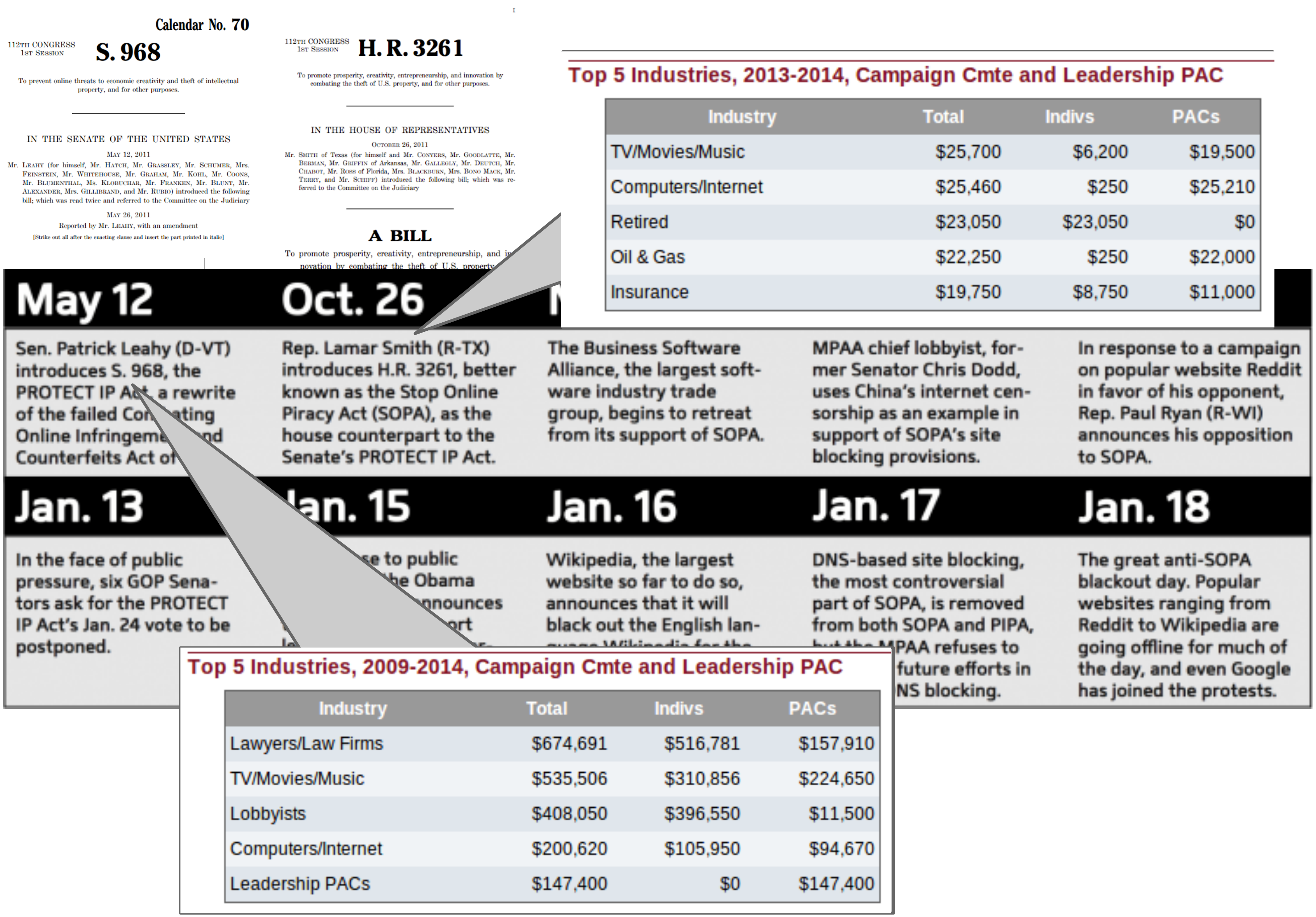
Mr. SMITH of Texas (for himself and Mr. CONYERS, Mr. GOODLATTE, Mr. BERMAN, Mr. GRIFFIN of Arkansas, Mr. GALLEGLY, Mr. DEUTCH, Mr. CHABOT, Mr. ROSS of Florida, Mrs. BLACKBURN, Mrs. BONO MACK, Mr. TERRY, and Mr. SCHIFF) introduced the following bill; which was referred to the Committee on the Judiciary

A BILL

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and

May 12	Oct. 26	Nov. 22	Dec. 8	Jan. 9
Sen. Patrick Leahy (D-VT) introduces S. 968, the PROTECT IP Act, a rewrite of the failed Combating Online Infringement and Counterfeits Act of 2010.	Rep. Lamar Smith (R-TX) introduces H.R. 3261, better known as the Stop Online Piracy Act (SOPA), as the house counterpart to the Senate's PROTECT IP Act.	The Business Software Alliance, the largest software industry trade group, begins to retreat from its support of SOPA.	MPAA chief lobbyist, former Senator Chris Dodd, uses China's internet censorship as an example in support of SOPA's site blocking provisions.	In response to a campaign on popular website Reddit in favor of his opponent, Rep. Paul Ryan (R-WI) announces his opposition to SOPA.

Jan. 13	Jan. 15	Jan. 16	Jan. 17	Jan. 18
In the face of public pressure, six GOP Senators ask for the PROTECT IP Act's Jan. 24 vote to be postponed.	In response to public opposition, the Obama administration announces that it will not support legislation that "...undermines the dynamic, innovative global internet."	Wikipedia, the largest website so far to do so, announces that it will black out the English language Wikipedia for the entirety of Wednesday, Jan. 18.	DNS-based site blocking, the most controversial part of SOPA, is removed from both SOPA and PIPA, but the MPAA refuses to renounce future efforts in favor of DNS blocking.	The great anti-SOPA blackout day. Popular websites ranging from Reddit to Wikipedia are going offline for much of the day, and even Google has joined the protests.



Calendar No. 70

112TH CONGRESS
1ST SESSION

S. 968

To prevent online threats to economic creativity and theft of intellectual property, and for other purposes.

IN THE SENATE OF THE UNITED STATES

MAY 12, 2011

Mr. LEAHY (for himself, Mr. HATCH, Mr. GRASSLEY, Mr. SCHUMER, Mrs. FEINSTEIN, Mr. WHITEHOUSE, Mr. GRAHAM, Mr. KOHL, Mr. COONS, Mr. BLUMENTHAL, Ms. KLOBUCHAR, Mr. FRANKEN, Mr. BLUNT, Mr. ALEXANDER, Mrs. GILLIBRAND, and Mr. RUBIO) introduced the following bill; which was read twice and referred to the Committee on the Judiciary

MAY 26, 2011

Reported by Mr. LEAHY, with an amendment

[Strike out all after the enacting clause and insert the part printed in italic]

I

112TH CONGRESS
1ST SESSION

H. R. 3261

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

OCTOBER 26, 2011

Mr. SMITH of Texas (for himself and Mr. CONYERS, Mr. GOODLATTE, Mr. BERMAN, Mr. GRIFFIN of Arkansas, Mr. GALLEGLY, Mr. DEUTCH, Mr. CHABOT, Mr. ROSS of Florida, Mrs. BLACKBURN, Mrs. BONO MACK, Mr. TERRY, and Mr. SCHIFF) introduced the following bill; which was referred to the Committee on the Judiciary

A BILL

To promote prosperity, creativity, entrepreneurship, and in novation by combating the theft of U.S. property

Top 5 Industries, 2013-2014, Campaign Cmte and Leadership PAC

Industry	Total	Indivs	PACs
TV/Movies/Music	$25,700	$6,200	$19,500
Computers/Internet	$25,460	$250	$25,210
Retired	$23,050	$23,050	$0
Oil & Gas	$22,250	$250	$22,000
Insurance	$19,750	$8,750	$11,000

May 12

Sen. Patrick Leahy (D-VT) introduces S. 968, the PROTECT IP Act, a rewrite of the failed Co ating Online Infringeme nd Counterfeits Act of

Oct. 26

Rep. Lamar Smith (R-TX) introduces H.R. 3261, better known as the Stop Online Piracy Act (SOPA), as the house counterpart to the Senate's PROTECT IP Act.

M

The Business Software Alliance, the largest software industry trade group, begins to retreat from its support of SOPA.

MPAA chief lobbyist, former Senator Chris Dodd, uses China's internet censorship as an example in support of SOPA's site blocking provisions.

In response to a campaign on popular website Reddit in favor of his opponent, Rep. Paul Ryan (R-WI) announces his opposition to SOPA.

Jan. 13

In the face of public pressure, six GOP Senators ask for the PROTECT IP Act's Jan. 24 vote to be postponed.

Jan. 15

se to public he Obama nnounces ort

Jan. 16

Wikipedia, the largest website so far to do so, announces that it will black out the English language Wikipedia for the

Jan. 17

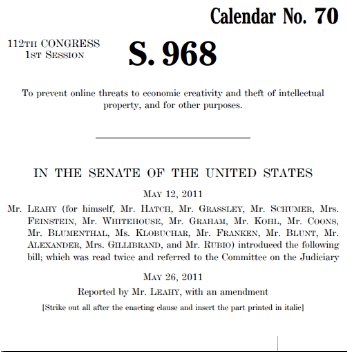
DNS-based site blocking, the most controversial part of SOPA, is removed from both SOPA and PIPA, but the MPAA refuses to future efforts in NS blocking.

Jan. 18

The great anti-SOPA blackout day. Popular websites ranging from Reddit to Wikipedia are going offline for much of the day, and even Google has joined the protests.

Top 5 Industries, 2009-2014, Campaign Cmte and Leadership PAC

Industry	Total	Indivs	PACs
Lawyers/Law Firms	$674,691	$516,781	$157,910
TV/Movies/Music	$535,506	$310,856	$224,650
Lobbyists	$408,050	$396,550	$11,500
Computers/Internet	$200,620	$105,950	$94,670
Leadership PACs	$147,400	$0	$147,400



Calendar No. 70

112TH CONGRESS
1ST SESSION

S. 968

To prevent online threats to economic creativity and theft of intellectual property, and for other purposes.

IN THE SENATE OF THE UNITED STATES

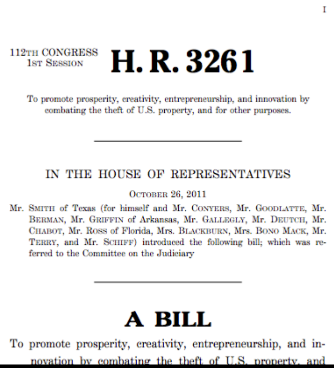
MAY 12, 2011

Mr. LEAHY (for himself, Mr. HATCH, Mr. GRASSLEY, Mr. SCHUMER, Mrs. FEINSTEIN, Mr. WHITEHOUSE, Mr. GRAHAM, Mr. KOHL, Mr. COONS, Mr. BLUMENTHAL, Ms. KLOBUCHAR, Mr. FRANKEN, Mr. BLUNT, Mr. ALEXANDER, Mrs. GILLIBRAND, and Mr. RUBIO) introduced the following bill; which was read twice and referred to the Committee on the Judiciary

MAY 26, 2011

Reported by Mr. LEAHY, with an amendment

[Strike out all after the enacting clause and insert the part printed in italic]



I

112TH CONGRESS
1ST SESSION

H. R. 3261

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

OCTOBER 26, 2011

Mr. SMITH of Texas (for himself and Mr. CONYERS, Mr. GOODLATTE, Mr. BERMAN, Mr. GRIFFIN of Arkansas, Mr. GALLEGLY, Mr. DEUTCH, Mr. CHABOT, Mr. ROSS of Florida, Mrs. BLACKBURN, Mrs. BONO MACK, Mr. TERRY, and Mr. SCHIFF) introduced the following bill; which was referred to the Committee on the Judiciary

A BILL

To promote prosperity, creativity, entrepreneurship, and innovation by combating the theft of U.S. property, and

May 12

Sen. Patrick Leahy (D-VT) introduces S. 968, the PROTECT IP Act, a rewrite of the failed Combating Online Infringement and Counterfeits Act of 2010.

Oct. 26

Rep. Lamar Smith (R-TX) introduces H.R. 3261, better known as the Stop Online Piracy Act (SOPA), as the house counterpart to the Senate's PROTECT IP Act.

Nov. 22

The Business Software Alliance, the largest software industry trade group, begins to retreat from its support of SOPA.

Dec. 8

MPAA chief lobbyist, former Senator Chris Dodd, uses China's internet censorship as an example in support of SOPA's site blocking provisions.

Jan. 9

In response to a campaign on popular website Reddit in favor of his opponent, Rep. Paul Ryan (R-WI) announces his opposition to SOPA.

Jan. 13

In the face of public pressure, six GOP Senators ask for the PROTECT IP Act's Jan. 24 vote to be postponed.

Jan. 15

In response to public opposition, the Obama administration announces that it will not support legislation that "...undermines the dynamic, innovative global internet."

Jan. 16

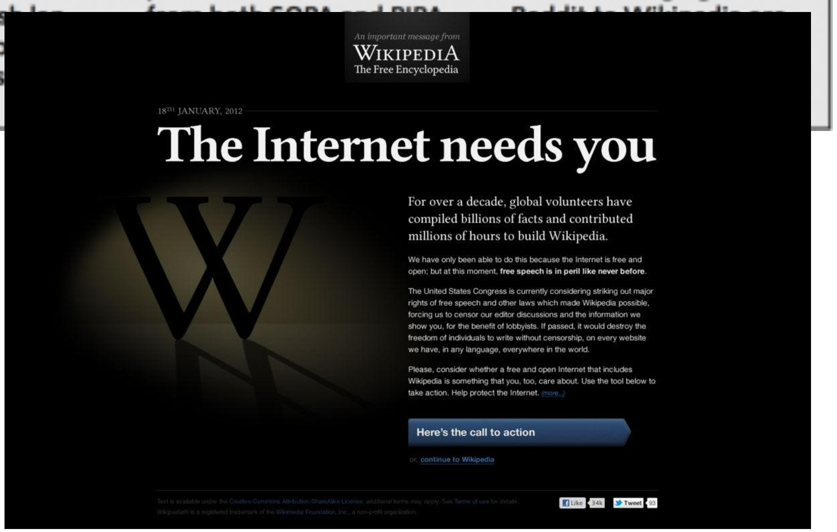
Wikipedia, the largest website so far to do so, announces that it will black out the Englis
guage Wikipedia fo
entirety of Wednes
Jan. 18.

Jan. 17

DNS-based site blocking, the most controversial part of SOPA, is removed
from both SOPA and PIPA

Jan. 18

The great anti-SOPA blackout day. Popular websites ranging from
Reddit to Wikipedia are



An important message from
WIKIPEDIA
The Free Encyclopedia

18TH JANUARY, 2012

The Internet needs you

For over a decade, global volunteers have compiled billions of facts and contributed millions of hours to build Wikipedia.

We have only been able to do this because the Internet is free and open; but at this moment, **free speech is in peril like never before**.

The United States Congress is currently considering striking out major rights of free speech and other laws which made Wikipedia possible, forcing us to censor our editor discussions and the information we show you, for the benefit of lobbyists. If passed, it would destroy the freedom of individuals to write without censorship, on every website we have, in any language, everywhere in the world.

Please, consider whether a free and open Internet that includes Wikipedia is something that you, too, care about. Use the tool below to take action. Help protect the Internet. (more...)

or, continue to Wikipedia

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of use for details.
Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.

Members of Congress's Positions on SOPA/PIPA, as tracked by ProPublica.org

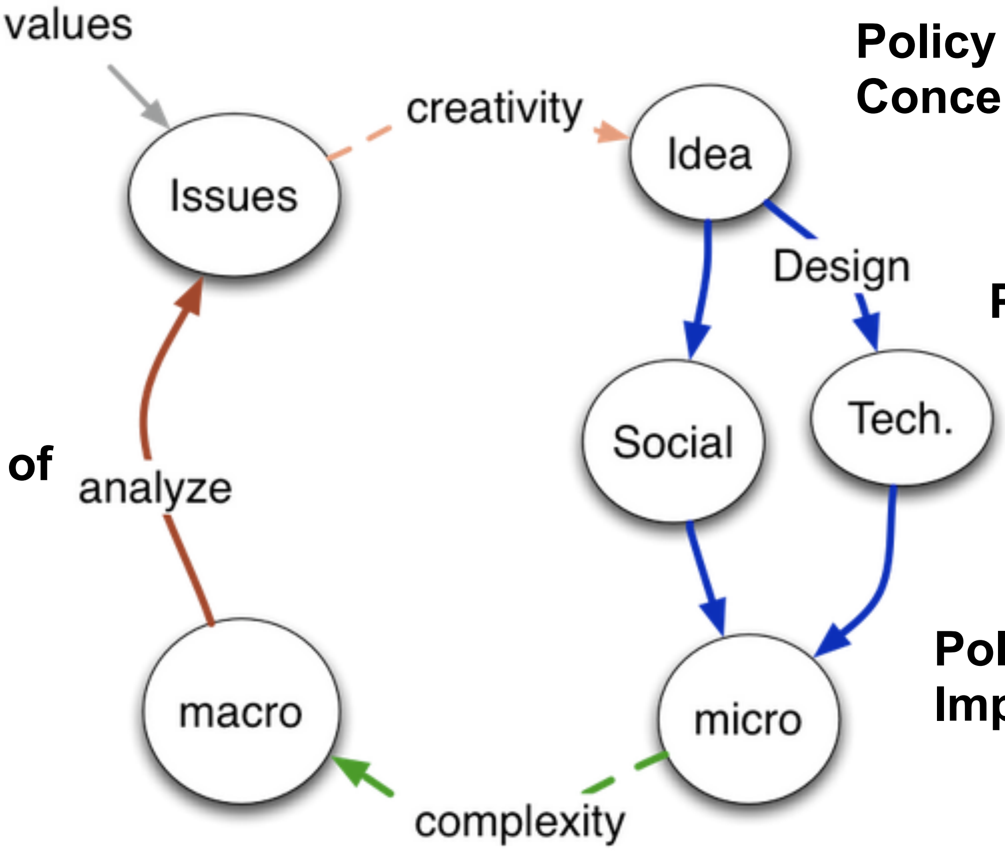
Jan. 18

80 supporters 31 opponents

Jan. 19

65 supporters 101 opponents

Web Science meets (Web) Governance?



Policy Conception

Policy Design

Analysis and understanding of policy implications

Policy Implementation

Review...

1. A Science of The Web
2. Web Architecture
3. Measuring the Web
4. The Web Science Method
5. Social Aspects of the Web
6. Web and other Governance

Assignment:

1. Preferential Attachment Simulator: http://bit.ly/18bd0p2
   - Try the THINGS TO TRY!
2. Excel-based Network Analysis Tutorial
   - Following instructions at: http://bit.ly/1a3mtzW
   - Install **NodeXL** from: http://bit.ly/1a3mnbx
   - Use **Senate 2007** data from: http://bit.ly/1a3mhAI
   - Play with other data at: http://bit.ly/1a3mJ1X
3. Social Network Exploration
   - Twitalizer: http://twitalyzer.com
   - TweetArchivist: http://tweetarchivist.com
   - MentionMap: http://mentionmapp.com
4. Create a Web Science Scenario:
   - Identify a (social) problem
   - Proposed an engineered solution
   - Identify how to measure, analyze, evaluate, iterate