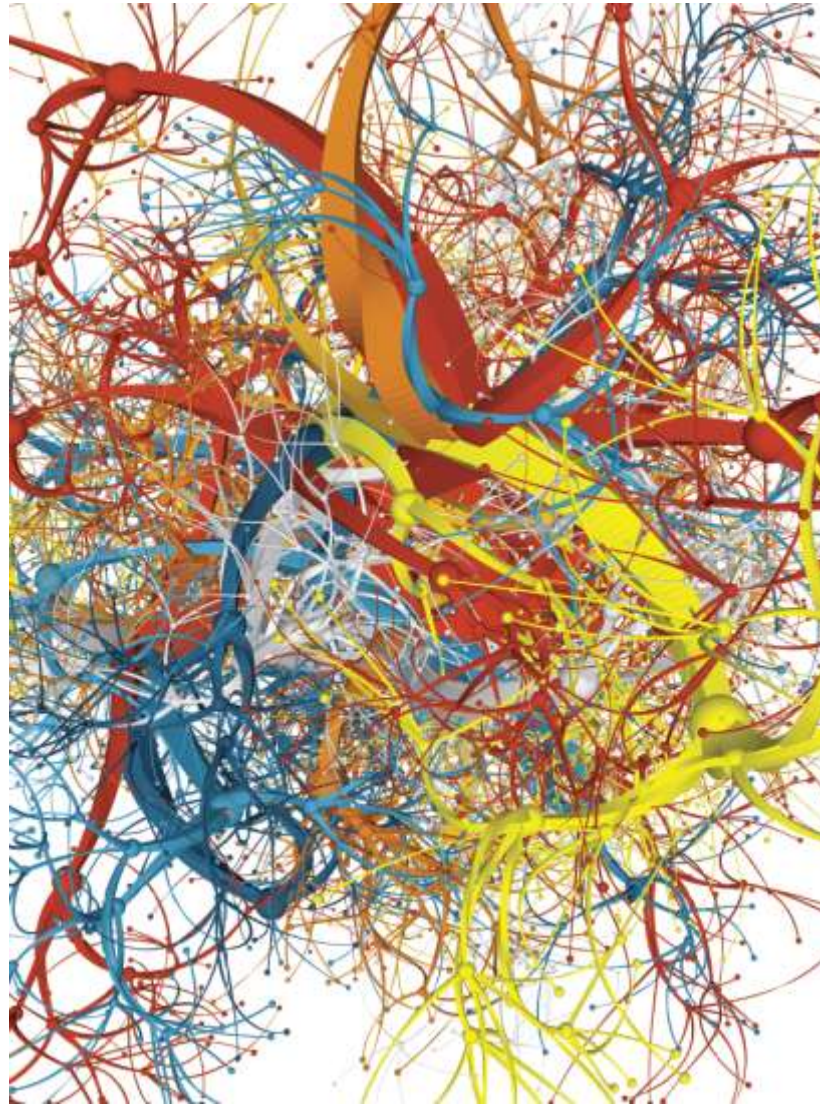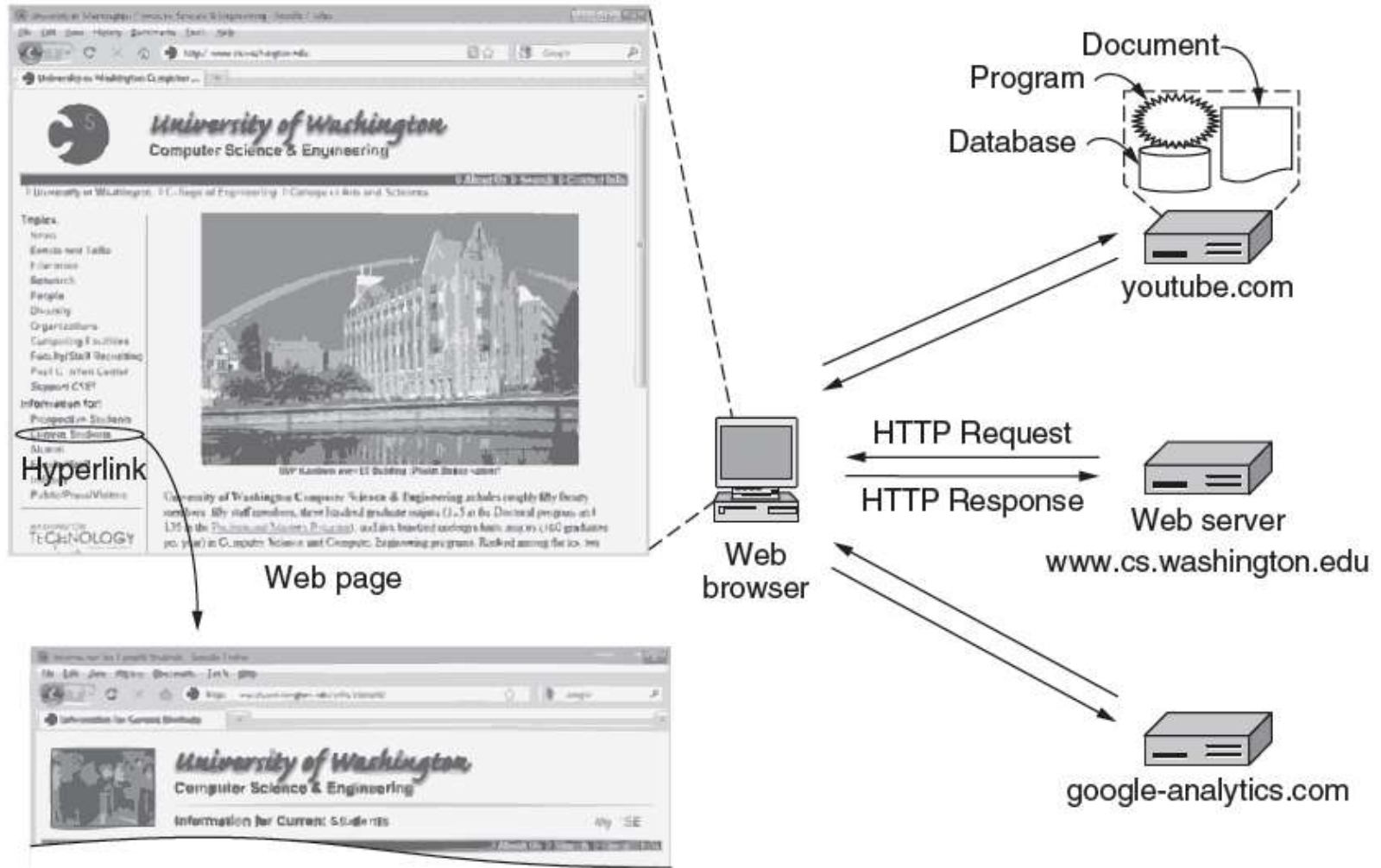# Web Science

Overview

- **The web**
  - History
  - Stats

- **"Web science"**
  - History
  - Suggested curriculum

- **What we're going to do**
  - Course details
  - Topics

A short history of the web

- **1989** Tim Berners-Lee at CERN
- **1990** HTTP/0.9, HTML, URLs, first text-based browser
- **1993** Marc Andreesen releases NCSA Mosaic, graphical browser
- **1993** CERN agrees to release protocol royalty-free
- **1994** Andreesen forms Netscape
- **1994** W3C formed, standardizing protocols, encouraging interoperability
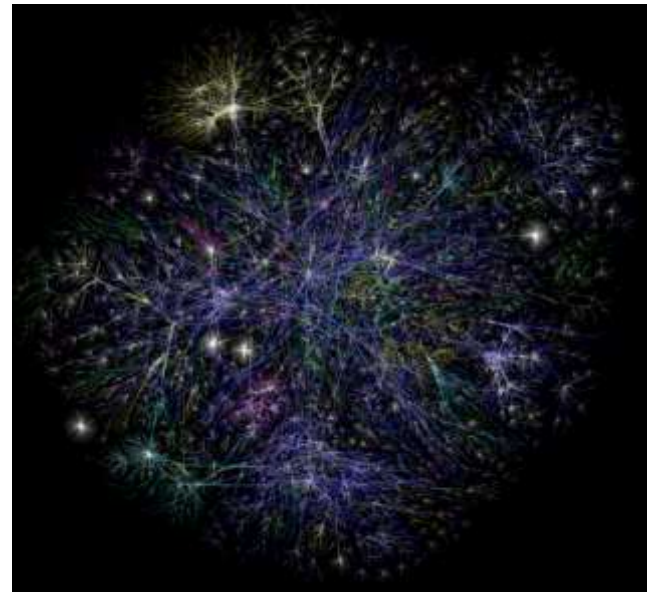
How does the web work?

What is the web?

- "The largest human information construct in history. The web is transforming society."
  - Web Science Trust

1989

2005

How big is the web?

- "We've known it for a long time: the web is big."
  - Google

Year	Unique URLs
1998	26 million
2000	1 billion
2004	8 billion
2008	1 trillion

- Human population, ~7 billion
  - 143 pages / person
- 1 billion Google search queries a day

http://funquisha.deviantart.com/art/ORIGAMI-ZEBRA-for-real-205044992

NEWS TECHNOLOGY

26 January 2011 Last updated at 12:59 ET

659 ≪ Share f ⦿ ✉ 🖨

Confusion over Egyptian blocks on web protest tools

Confusion surrounds the use of web tools such as Twitter and Facebook that have been used by protesters in Egypt to co-ordinate action.

The Egyptian government denied taking any action to restrict use of the web, saying it respected freedom of expression.

However, Twitter said it is being blocked but said many people have found ways round the restrictions.

Egyptian anti-government activists clashed with police in defiance of a ban on protests

8

26 January 2011 Last updated at 05:39 ET

1.3K Share f 🐦 ✉ 🖨

Twitter and web video site face clampdown in Egypt

By Jonathan Fildes
Technology reporter, BBC News

Egypt appears to have clamped down on web services, such as Twitter, that have been used to help organise anti-government protests in Cairo.

Twitter confirmed that its service has been blocked in Egypt on Tuesday from around 1600GMT.

A Swedish mobile video site called Bambuser also reported that it had been blocked around the same time.

REUTERS

Riot police used water cannons to disperse demonstrators

Business Applications	Comms & Networking	Data Centre & IT Infrastructure	Information Management

News / IT Case studies / Perspectives & Trends / Features / Company Analys

HOME > COMMS AND NETWORKING > NEWS

De-risking the cloud with end-to-end SLAS
Smart Wired Cloud Seminar

Egypt blocks Internet access amid protests

28 JANUARY 2011 Daniel Shane

AA

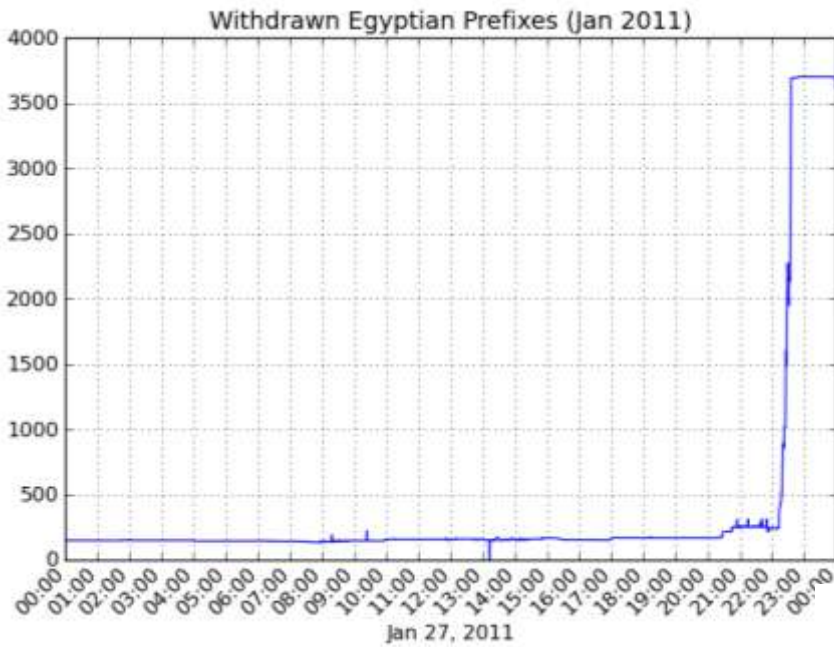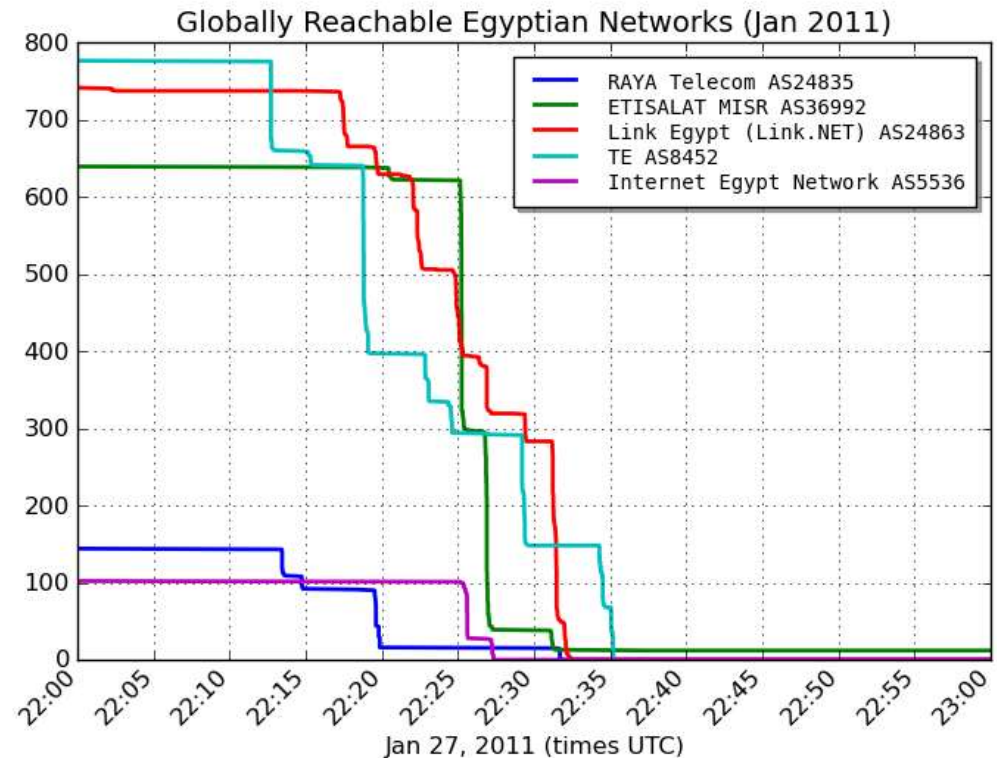**Government orders telcos to block web access as protestors take to the streets**

The Egyptian government has called on telecommunications providers in the country to block access to the Internet in response to widespread civil unrest.

Vodafone Egypt, one of the largest operators in the country not controlled by the state, today said it has disabled access following pressure from authorities.

Withdrawn Egyptian Prefixes (Jan 2011)

"What happens when you disconnect a modern economy and 80,000,000 people from the Internet?"

http://www.renesys.com/blog/2011/01/egypt-leaves-the-internet.shtml

Globally Reachable Egyptian Networks (Jan 2011)

RAYA Telecom AS24835
ETISALAT MISR AS36992
Link Egypt (Link.NET) AS24863
TE AS8452
Internet Egypt Network AS5536

Demonstrators in Cairo's Tahrir Square, February 8, 2011

What is "web science"?

- ## Science 2006
  - MIT, University of Southhampton

Creating a Science of the Web

Tim Berners–Lee[1], Wendy Hall[2], James Hendler[3], Nigel Shadbolt[2], Daniel J. Weitzner[1]

[1]Computer Science and Artificial Intelligence Laboratory, MIT, USA.
[2]School of Electronics and Computer Science, University of Southampton, UK.
[3]Computer Science Department, University of Maryland, USA.

Since its inception, the World Wide Web has changed the ways scientists communicate, collaborate, and educate. There is, however, a growing realization among many researchers that a clear research agenda aimed at understanding the current, evolving, and potential Web is needed. If we want to model the Web; if we want to understand the architectural principles that have provided for its growth; and if we want to be sure that it supports the basic social values of trustworthiness, privacy, and respect for social boundaries, then we must chart out a research agenda that targets the Web as a primary focus of attention.

What is "web science"?

- "Web science...is inherently interdisciplinary and integrates computer and information sciences with a multitude of disciplines including sociology, economics, political science, law, management, language and communication, geography and psychology. "

Web science curriculum topics

History of the Web

Forerunners (Otlet, Wells, Bush, Engelbart, Nelson) - information systems, concepts, early computer systems

Hypertext Community - information systems

Internet history - DARPA, IP, TCP, FTP, WAIS, GOPHER

W3C History - See W3C timeline

Building the Web

Web Architecture (HTTP, HTML, URI, XML, XSLT, JavaScript, AJAX)

Key Algorithms

Community Inclusion- Incentives for Innovation - Openness / universality

Decentralisation

Governance

Standards

Web science curriculum topics

The Web in Society

E-commerce

IP / copyright

Privacy

Co-evolution of society and web

Culture and technology

Systems theory

Social structures and processes

Groups and identity

Commercial structures and economics

Globalisation

Social capital and power inequality

Collective intelligence

Web science curriculum topics

Deploying the Web - Operationalising Web Science for a World of International Commerce

- Business Strategy
- Information systems (basics of)
- Cloud computing infrastructure
- Policy
- Regulation and security
- Sector-specific info
- Online markets
- Design vs evolution
- International context - developed and developing world
- Profit vs common good
- Software / hardware context (speeds etc)

Web science curriculum topics

Analysing the Web

Methodologies (build around case studies)

Uncertainties and critical thinking

Graph theory

Power laws

Statistics / regression analysis

Networks - game theory, social network analysis, ANT

Web mining

Understanding Web Users

Surveys

Qualitative

This course

- Focus on technical side
  - Learn interesting technologies
  - Build stuff

- Objectives:
  - Understand and the resources, data formats, and protocols used in the Internet
  - Build robust and load balanced client server software
  - Understand and use public/private key cryptography
  - Understand state management on the Internet

Tentative topics

- Data formats
  - ASCII/Unicode
  - Audio, images, video
- Compression
  - The only thing making Netflix/YouTube work on today's networks
- Encryption
  - Symmetric encryption, e.g. DES, 3DES, AES
  - Public/private encryption, e.g. SSL
  - Authentication

Tentative topics

- State management
  - Storing state at the client or server
  - Cookies, session variables, query string, databases
- Technologies
  - Client side, server side scripting
  - Apache, PHP, MySQL, OpenSSL
  - HTML5, Adobe Flex
- Web search
  - PageRank algorithm
- Web services
  - SOAP, REST

Tentative topics

- Crowdsourcing
  - Amazon Mechanical Turk
- Availability
  - Providing fault tolerance
  - 99.999%
- Scalability
  - Scaling to million/billions of hits