

(CMS/CS/EE 144)

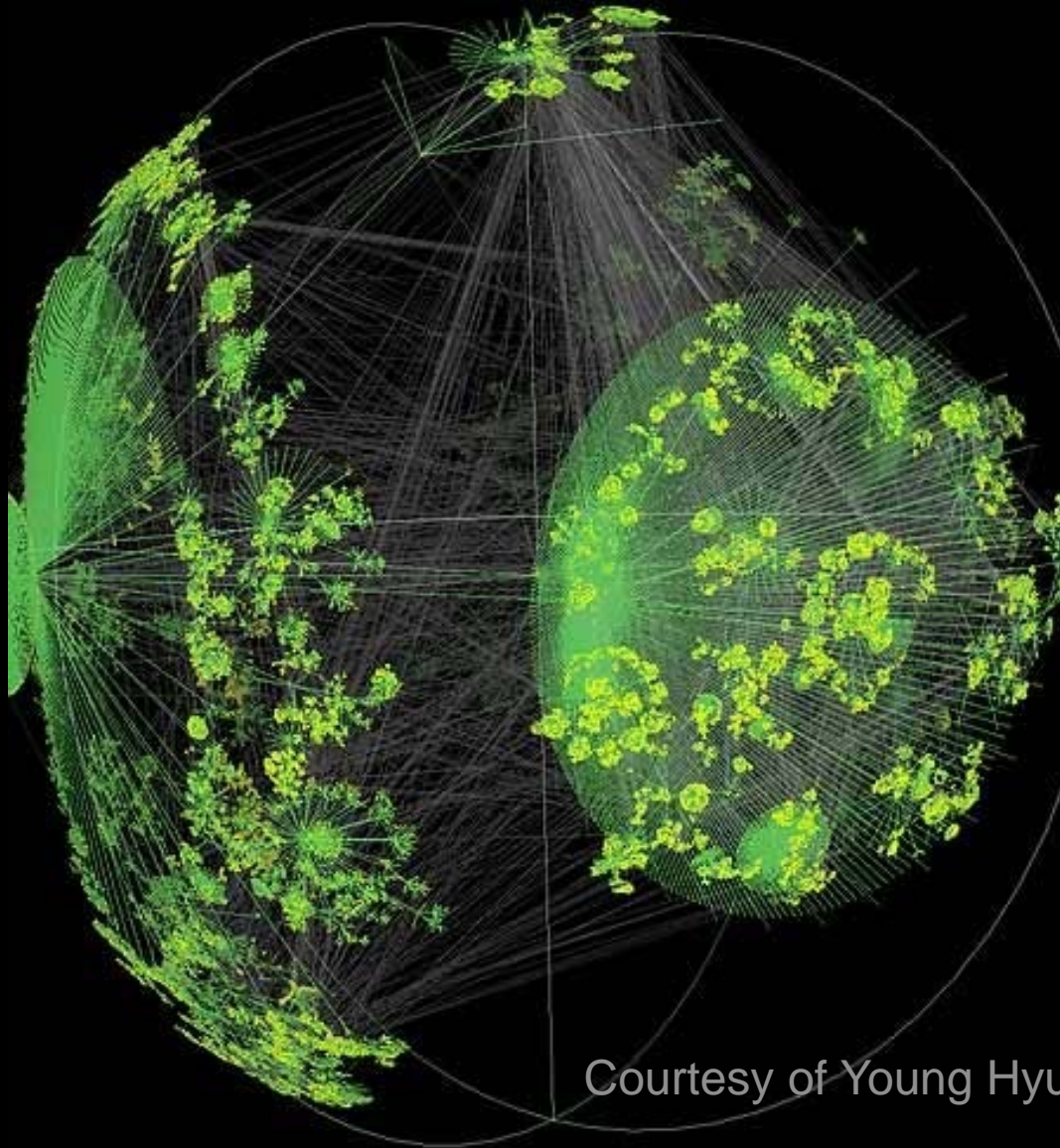
Networks: Structure & Economics

>Adam Wierman

What is this course about?
& what will it require of you?

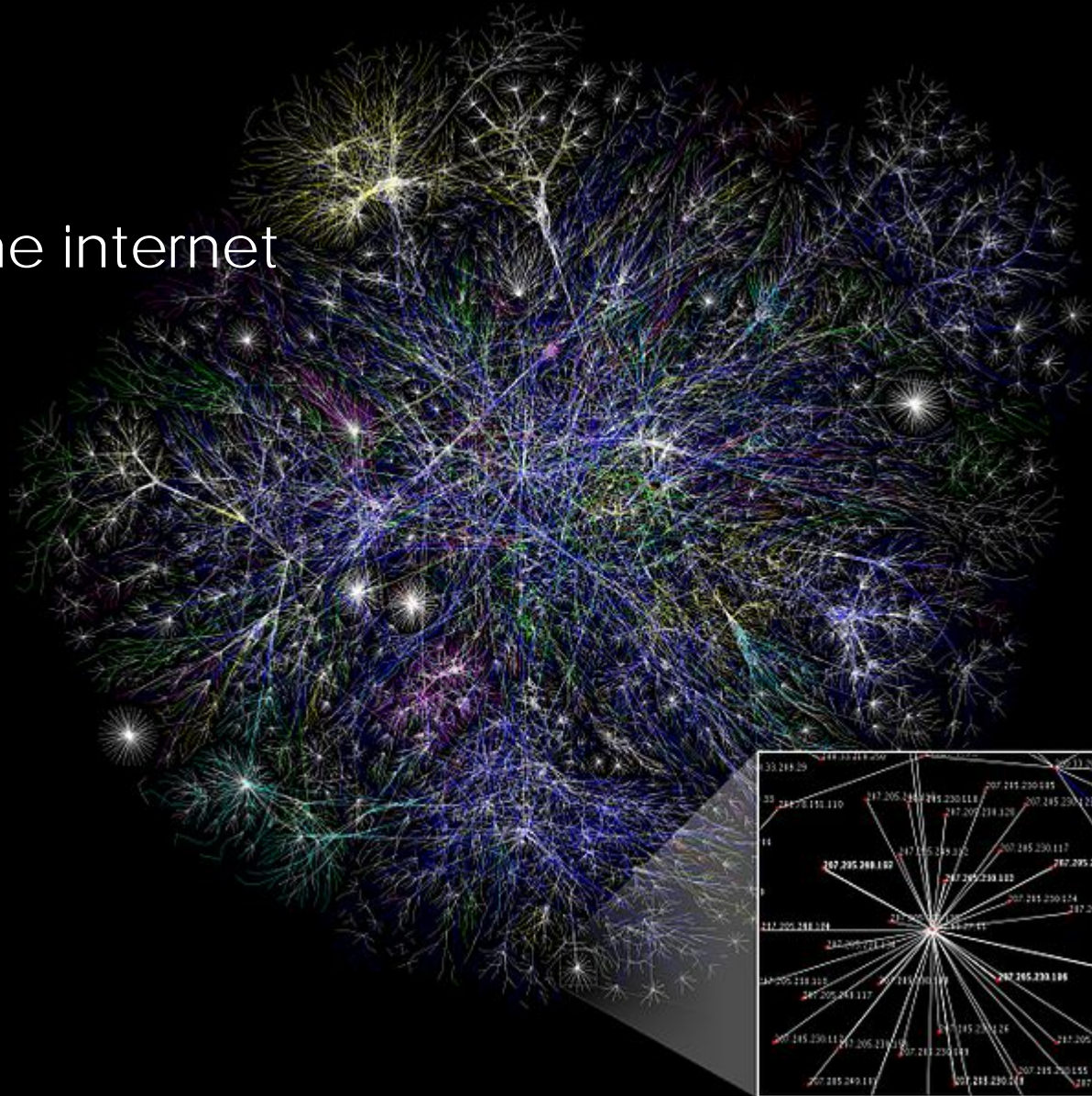
This course is about what the networks
we interact with “look” like
... how they “work”
... and how to exploit them!

The web



Courtesy of Young Hyun, CAIDA

The internet





facebook

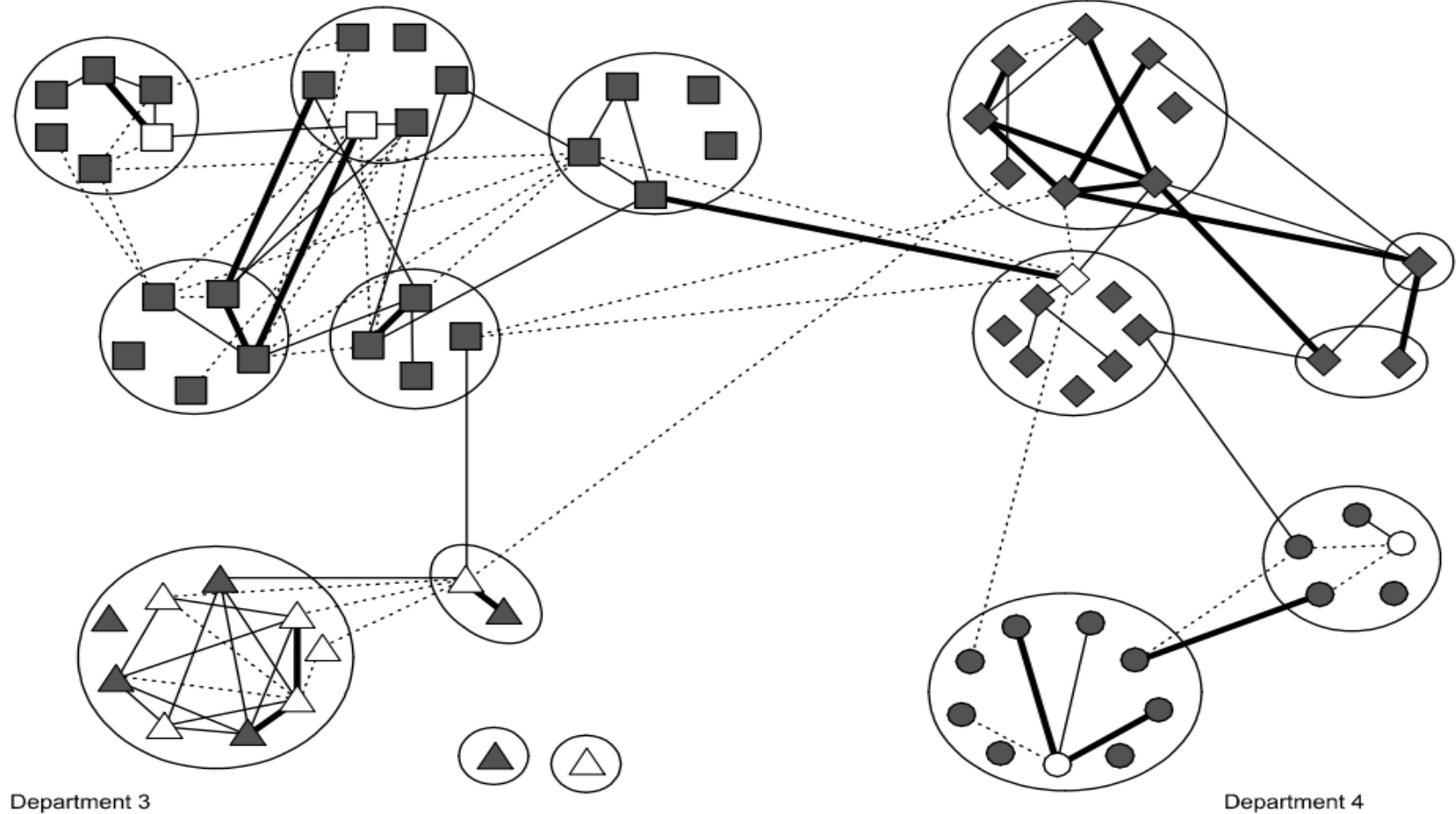
December 2010



twitter

Department 1

Department 2

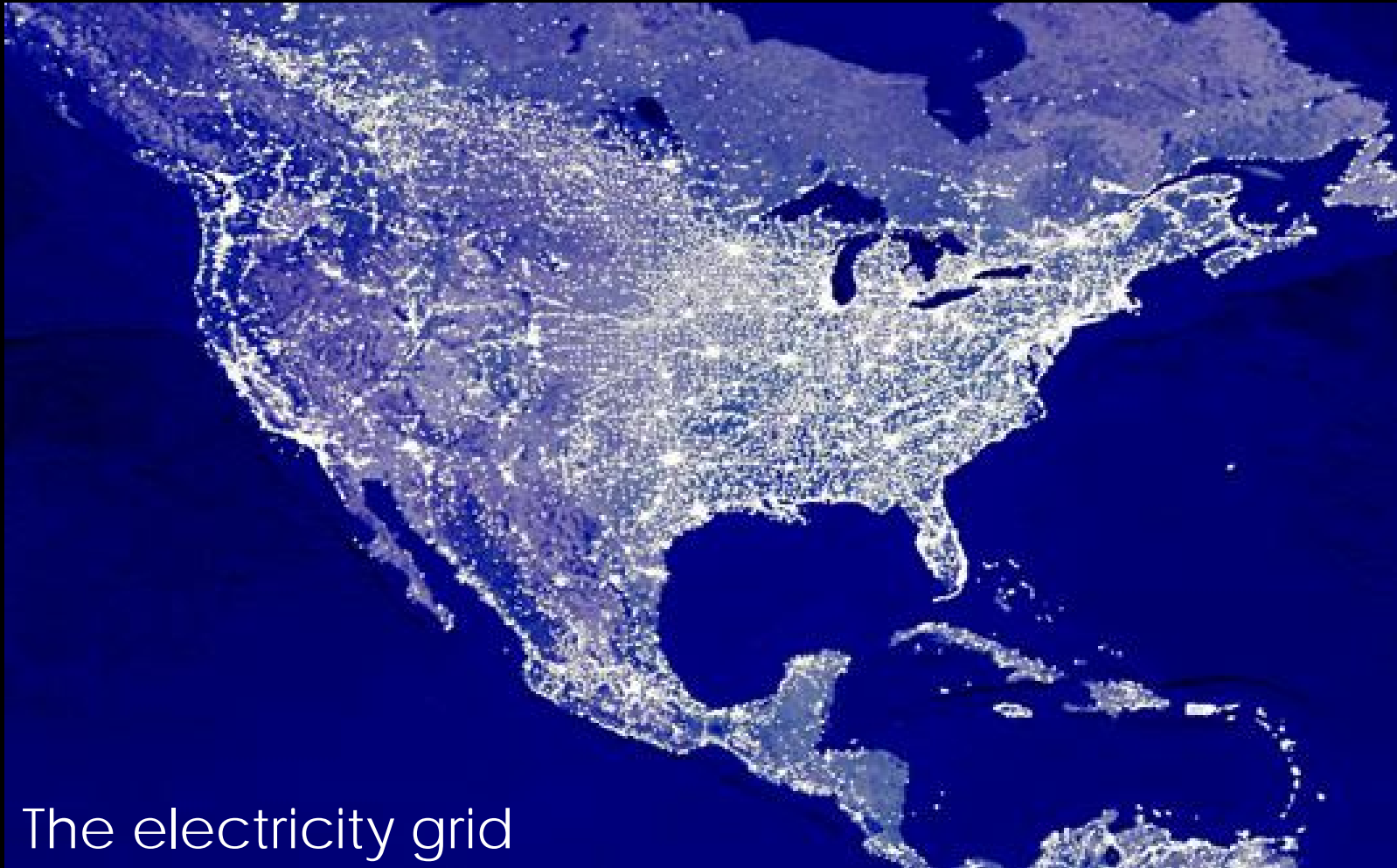


Department 3

Department 4

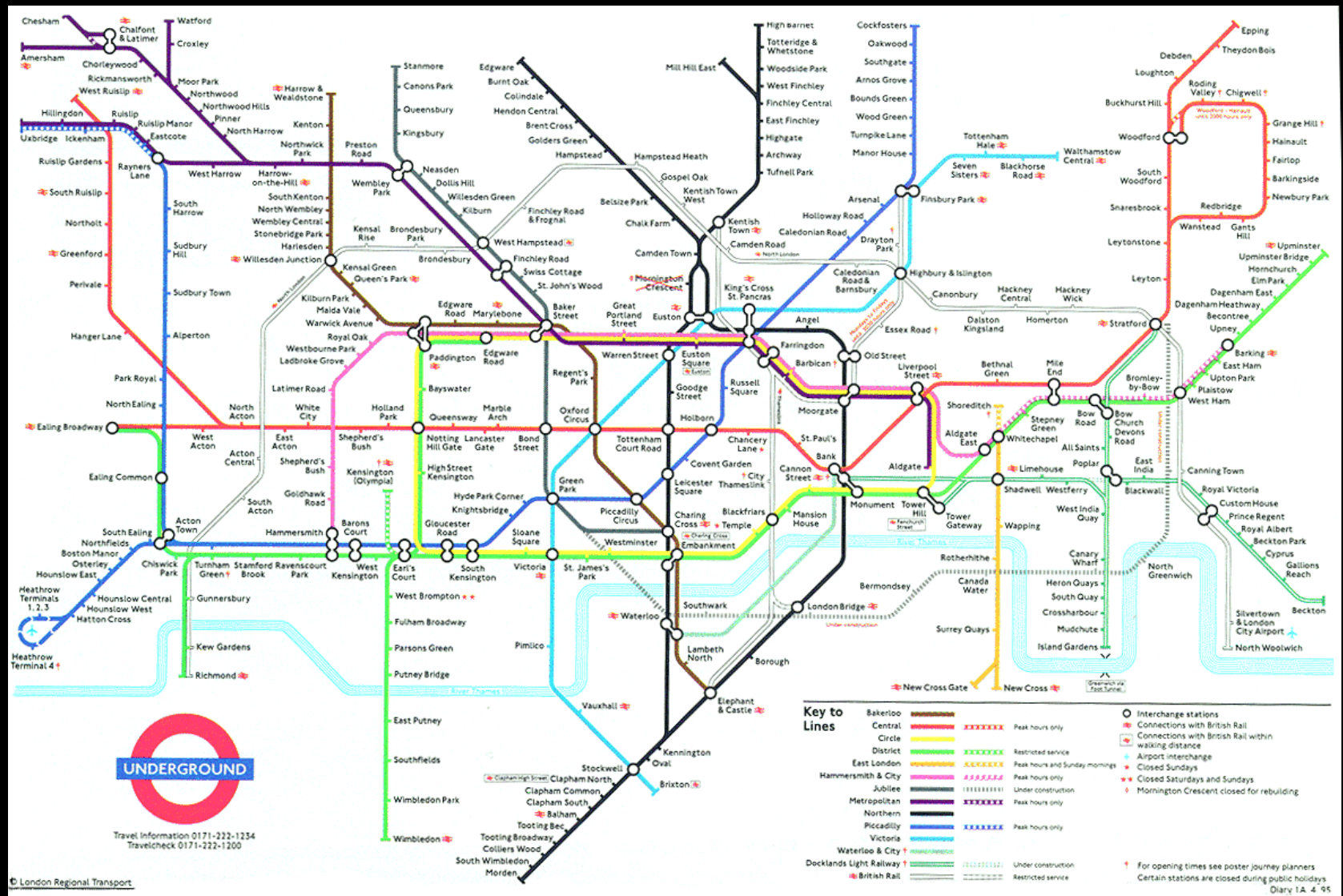


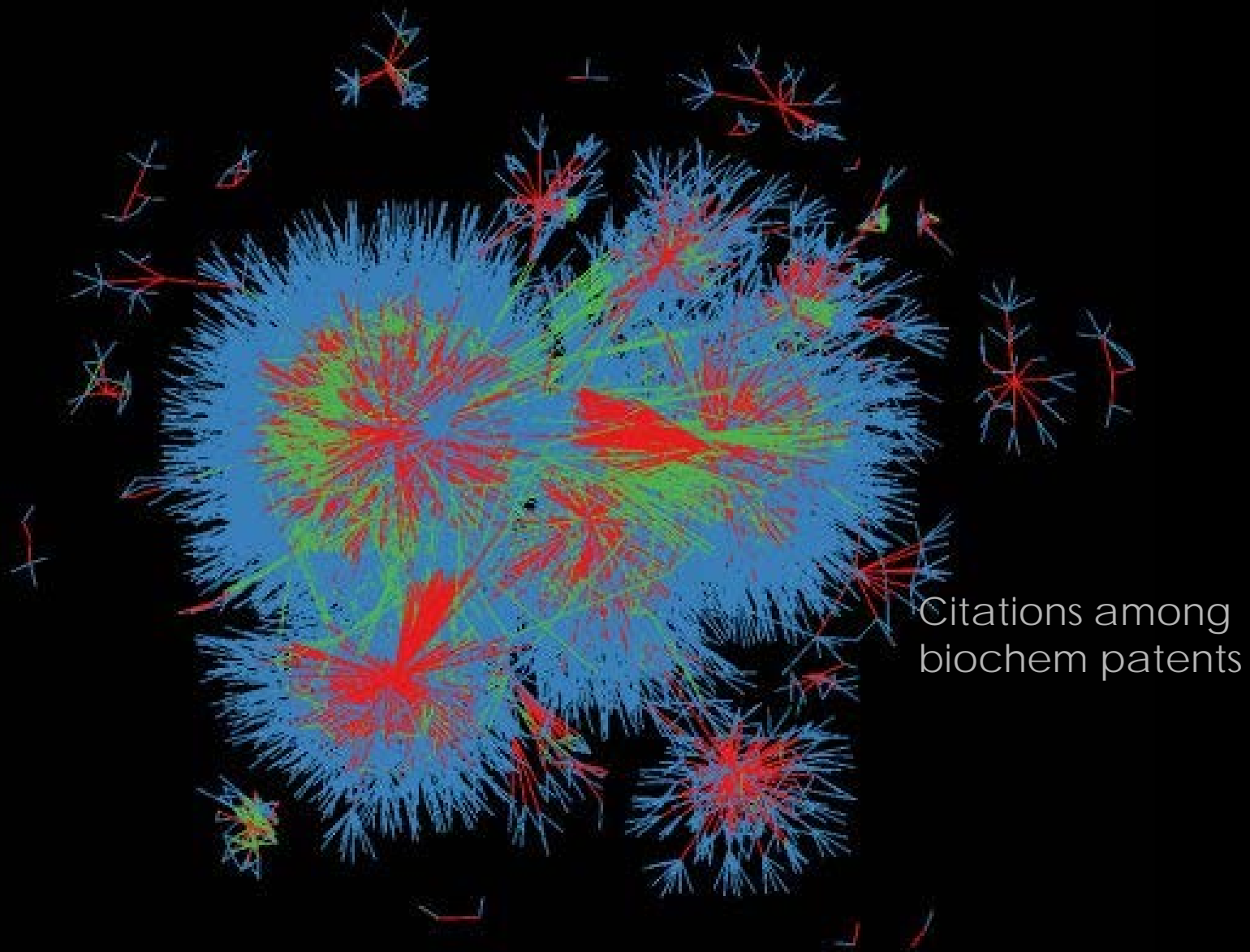
(from Leeat Yariv)

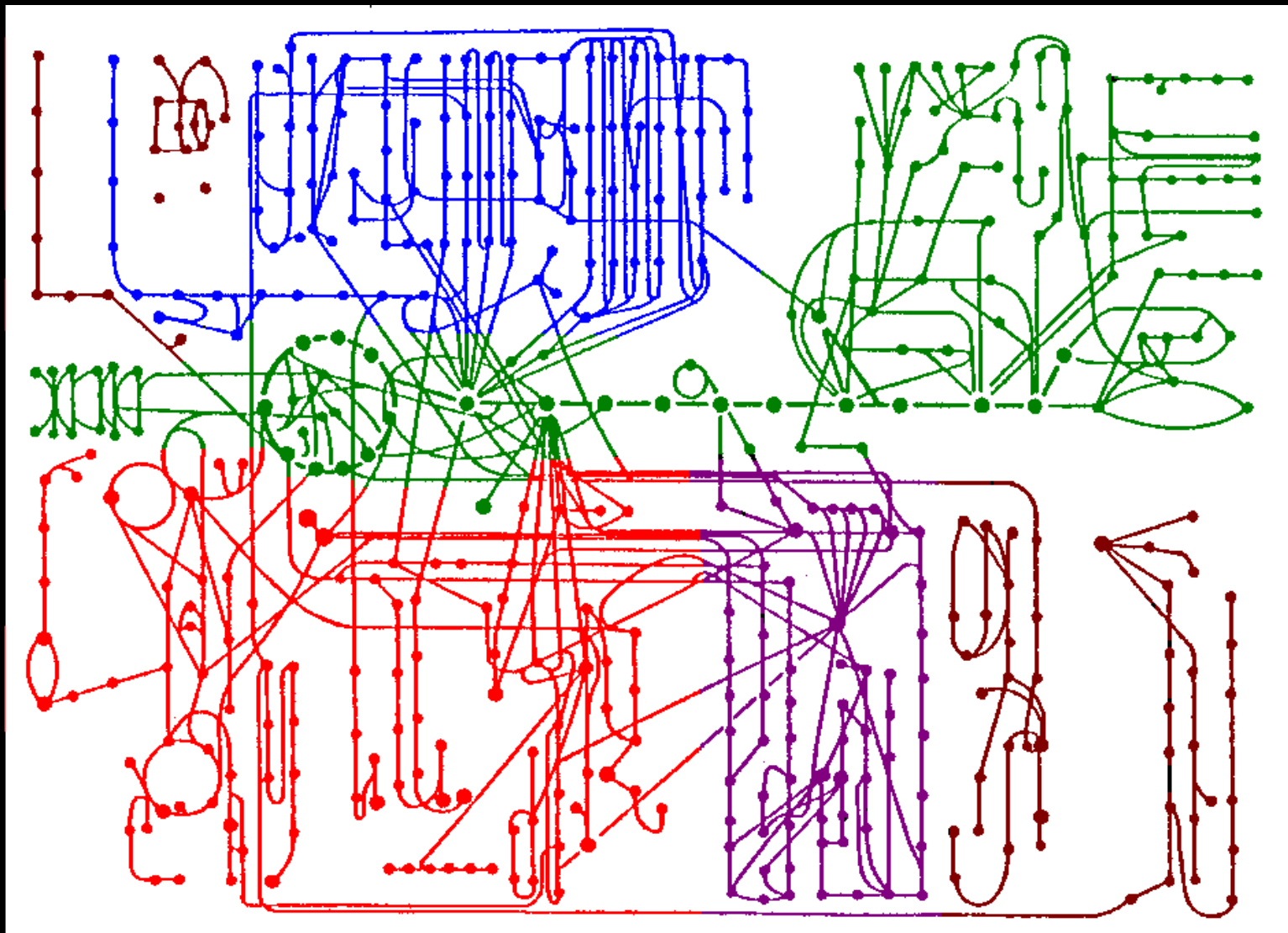


The electricity grid

Transportation networks







Biochemical
pathways of
malaria

What structural properties do networks have?
...what causes these properties to emerge?
...why do many networks look so similar?


These properties underlie the design of
many modern tech companies & startups!

You use Google, Facebook, etc every day,
BUT you probably know very little about
how they work...

...these are the companies you want a job from!


...the ideas behind these companies are where
the modern networking research problems are!

How does google rank search results?

Google 

[Web](#) [Images](#) [Maps](#) [Shopping](#) [News](#) [More ▾](#) [Search tools](#)

About 742,000,000 results (0.31 seconds)

Ads related to tablet 

[iPad mini - Every inch an iPad - apple.com](#)
www.apple.com/
Learn more.
Buy now at Apple Online Store - Beautiful 7.9 inch display - Over 275,000 apps

[Tablet Hotels™ - Great Deals on Boutique Hotels](#)
www.tablethotels.com/
Best Rate Guarantee Reserve Online
295 people +1'd this page
Reserve A Hotel - Great Hotel Deals - Private Sales - New Hotels


[New Dell Tablets - dell.com](#)
www.dell.com/Windows8 - ★★★★★ 3,139 seller reviews
Check Out Dell's New Line of Tablets w/ Intel® Core™ Today!









[Tablets and Tablet PC Reviews - CNET Reviews](#)
reviews.cnet.com/tablets/
The best version of the iPad yet is also the best tablet on the market. ... The 1.6GHz Tegra 3 and DDR3 RAM make this the fastest Android tablet available. Drop.
[Best 5 tablets - Best Android tablets - Tablets buying guide - Samsung Tablets](#)

[Tablets: Android Tablet, Windows Tablet PC, Tablet Computers ...](#)
www.bestbuy.com/site/iPad-Tablets/Tablets/pomcat209000050008.c?..
Shop for a Tablet PC from Best Buy and select from Android Tablets, ... Buy a tablet, iPad®, laptop or desktop and get a one-year membership for only \$99.99.
[Tablets: Page 2 - Android - Samsung - Samsung Galaxy Tab 2 7.0](#)

[Tablets & Tablet PCs: Compare, Buy, and Review on Amazon.com](#)
www.amazon.com/b?ie=UTF8&node=1232597011
Read actual customer reviews and ratings on tablets and tablet PCs. Find low prices on a wide selection of tablets.


[Tablet Reviews | Tablets Review | PCMag.com](#)
www.pcmag.com > [Product Guides](#) > [Laptops & Notebooks](#)
PC Magazine provides up-to-date coverage and product reviews of tablets.

[Shop for tablet on Google](#) Sponsored 

 Sony 16Gb Xperia Tablet S \$399.99 Sony - Official..	 SAMSUNG Galaxy Tab 2... \$199.00 Newegg.c...	 Mach Speed Trio Eclipse... \$69.99 Sears	 Motorola Xoom 32GB WiFi A... \$229.99 N1 wireless
 Samsung Galaxy Tab 2... \$249.99 Medical Arts...	 Google Nexus 7 Black Wi-Fi... \$249.00 Walmart	 Samsung 8GB Galaxy Tab 2... \$199.99 Abt Electr...	 Maylong 7" Android Tabl... \$129.99 Target

Shop by brand

[Apple](#) [ASUS](#) [Samsung](#) [Archos](#) [Acer](#)

Ads 

[Google Nexus 7](#)
play.google.com/
The 7" tablet with mobile data.
Thin, light, portable. Learn more.
158,058 people +1'd or follow Nexus

[Amazon Tablet - Only \\$159](#)
www.amazon.com/kindlefire
★★★★★ 1,575 seller reviews
All-New Kindle Fire.
Web, Movies, Apps, Reading & More.

How does google place ads?

Google

tablet

Web Images Maps Shopping News More Search tools

About 742,000,000 results (0.31 seconds)

Ads related to tablet

iPad mini - Every inch an iPad - apple.com
www.apple.com/
Learn more.
Buy now at Apple Online Store - Beautiful 7.9 inch display - Over 275,000 apps

Tablet Hotels™ - Great Deals on Boutique Hotels
www.tablethotels.com/
Best Rate Guarantee Reserve Online
295 people +1'd this page
Reserve A Hotel - Great Hotel Deals - Private Sales - New Hotels

New Dell Tablets - dell.com
www.dell.com/Windows8 - ★★★★★ 3,139 seller reviews
Check Out Dell's New Line of Tablets w/ Intel® Core™ Today!









Tablets and Tablet PC Reviews - CNET Reviews
reviews.cnet.com/tablets/
The best version of the iPad yet is also the best tablet on the market. ... The 1.6GHz Tegra 3 and DDR3 RAM make this the fastest Android tablet available. Drop.
Best 5 tablets - Best Android tablets - Tablets buying guide - Samsung Tablets

Tablets: Android Tablet, Windows Tablet PC, Tablet Computers ...
www.bestbuy.com/site/iPad-Tablets/Tablets/pomcat209000050008.c?..
Shop for a Tablet PC from Best Buy and select from Android Tablets, ... Buy a tablet, iPad®, laptop or desktop and get a one-year membership for only \$99.99.
Tablets: Page 2 - Android - Samsung - Samsung Galaxy Tab 2 7.0

Tablets & Tablet PCs: Compare, Buy, and Review on Amazon.com
www.amazon.com/b?ie=UTF8&node=1232597011
Read actual customer reviews and ratings on tablets and tablet PCs. Find low prices on a wide selection of tablets.

Tablet Reviews | Tablets Review | PCMag.com
www.pcmag.com > Product Guides > Laptops & Notebooks
PC Magazine provides up-to-date coverage and product reviews of tablets.

Shop for tablet on Google Sponsored

 Sony 16Gb Xperia Tablet S \$399.99 Sony - Official..	 SAMSUNG Galaxy Tab 2... \$199.00 Newegg.c...	 Mach Speed Trio Eclipse... \$69.99 Sears	 Motorola Xoom 32GB WiFi A... \$229.99 N1 wireless
 Samsung Galaxy Tab 2... \$249.99 Medical Arts...	 Google Nexus 7 Black Wi-Fi... \$249.00 Walmart	 Samsung 8GB Galaxy Tab 2... \$199.99 Abt Electr...	 Maylong 7" Android Tab... \$129.99 Target

Shop by brand


[Apple](#) [ASUS](#) [Samsung](#) [Archos](#) [Acer](#)

Ads

Google Nexus 7
play.google.com/
The 7" tablet with mobile data. Thin, light, portable. Learn more.
158,058 people +1'd or follow Nexus


Amazon Tablet - Only \$159
www.amazon.com/kindlefire
★★★★★ 1,575 seller reviews
All-New Kindle Fire.
Web, Movies, Apps, Reading & More.

How does google serve the content so quickly?

Google 

[Web](#) [Images](#) [Maps](#) [Shopping](#) [News](#) [More ▾](#) [Search tools](#)

About 742,000,000 results (0.31 seconds)

Ads related to tablet 

[iPad mini - Every inch an iPad - apple.com](#)
www.apple.com/
Learn more.
Buy now at Apple Online Store - Beautiful 7.9 inch display - Over 275,000 apps

[Tablet Hotels™ - Great Deals on Boutique Hotels](#)
www.tablethotels.com/
Best Rate Guarantee Reserve Online
295 people +1'd this page
Reserve A Hotel - Great Hotel Deals - Private Sales - New Hotels


[New Dell Tablets - dell.com](#)
www.dell.com/Windows8 - ★★★★★ 3,139 seller reviews
Check Out Dell's New Line of Tablets w/ Intel® Core™ Today!









[Tablets and Tablet PC Reviews - CNET Reviews](#)
reviews.cnet.com/tablets/
The best version of the iPad yet is also the best tablet on the market. ... The 1.6GHz Tegra 3 and DDR3 RAM make this the fastest Android tablet available. Drop.
[Best 5 tablets - Best Android tablets - Tablets buying guide - Samsung Tablets](#)

[Tablets: Android Tablet, Windows Tablet PC, Tablet Computers ...](#)
www.bestbuy.com/site/iPad-Tablets/Tablets/pomcat209000050008.c?...
Shop for a Tablet PC from Best Buy and select from Android Tablets, ... Buy a tablet, iPad®, laptop or desktop and get a one-year membership for only \$99.99.
[Tablets: Page 2 - Android - Samsung - Samsung Galaxy Tab 2 7.0](#)

[Tablets & Tablet PCs: Compare, Buy, and Review on Amazon.com](#)
www.amazon.com/b?ie=UTF8&node=1232597011
Read actual customer reviews and ratings on tablets and tablet PCs. Find low prices on a wide selection of tablets.


[Tablet Reviews | Tablets Review | PCMag.com](#)
www.pcmag.com > [Product Guides](#) > [Laptops & Notebooks](#)
PC Magazine provides up-to-date coverage and product reviews of tablets.

[Shop for tablet on Google](#) Sponsored 

 Sony 16Gb Xperia Tablet S \$399.99 Sony - Official..	 SAMSUNG Galaxy Tab 2... \$199.00 Newegg.c...	 Mach Speed Trio Eclipse... \$69.99 Sears	 Motorola Xoom 32GB WiFi A... \$229.99 N1 wireless
 Samsung Galaxy Tab 2... \$249.99 Medical Arts...	 Google Nexus 7 Black Wi-Fi... \$249.00 Walmart	 Samsung 8GB Galaxy Tab 2... \$199.99 Abt Electr...	 Maylong 7" Android Tab... \$129.99 Target

Shop by brand

[Apple](#) [ASUS](#) [Samsung](#) [Archos](#) [Acer](#)

Ads 

[Google Nexus 7](#)
play.google.com/
The 7" tablet with mobile data.
Thin, light, portable. Learn more.
158,058 people +1'd or follow Nexus

[Amazon Tablet - Only \\$159](#)
www.amazon.com/kindlefire
★★★★★ 1,575 seller reviews
All-New Kindle Fire.
Web, Movies, Apps, Reading & More.

What goes into data centers?
How do we build applications for them?

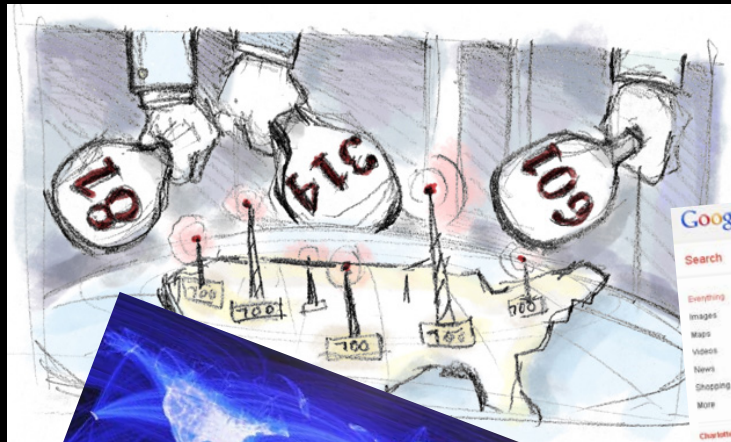


Google's Oregon data center
A 1.2 million core warehouse-sized computer

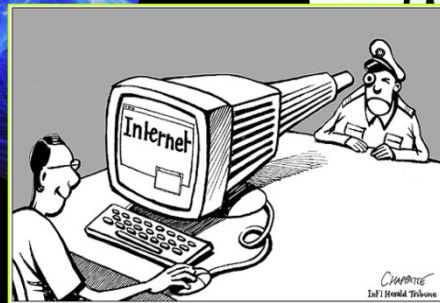
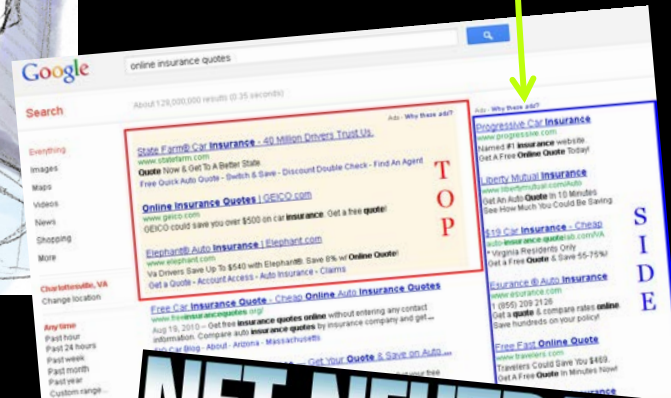
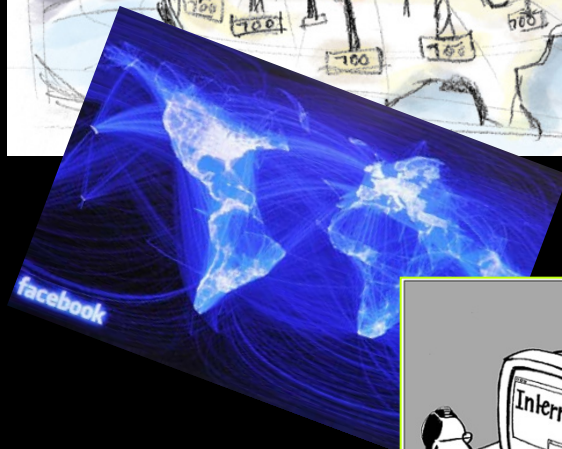
How does Facebook exploit its knowledge of the social network?

The screenshot shows a Facebook profile page for Adam Wierman. The top navigation bar includes the Facebook logo, a search bar, and the user's name 'Adam Wierman' with links to 'Home', a lock icon, and a settings gear. The left sidebar contains navigation links: 'FAVORITES' (News Feed, Messages, Events, Photos), 'ADS' (Ads Manager), 'PAGES' (The ideas behind the Web, Pages Feed, Like Pages), 'GROUPS' (Next Year's Our Year, Carnegie Mellon Alumni, Carnegie Mellon Los..., IEEE STC on Sustain..., Create Group...), 'APPS' (App Center, Games Feed, Music, Notes, Links, Pokes, Course Finder), and 'MORE'. The main content area shows the 'Update Status' section with a text input field containing 'How are you doing, Adam?'. Below this is a post from Fred Zhao, who likes 'Professor Dalves's School of Pedantry'. The post includes a photo of a man and the text 'Professor Dalves's School of Pedantry Teacher 222 likes'. Below the post is a link shared by Frank Pfenning, titled 'Press Release: Frank Pfenning Named Head of Carnegie Mellon's Computer Science Department -Carnegie.' with a URL 'www.cmu.edu' and contact information 'Contact: Byron Spice / 412-268-9068 /'. The right sidebar contains several recommendations: 'Alan Rutledge's birthday is today', 'Create Event', '1 request from Linda Höglund', 'Sponsored' (Create an Ad), 'Fred Zhao likes Professor Dalves's School of Pedantry.', 'Professor Dalves's School of Pedantry' (Like), 'Do You Dream in Code?' (Get a cool job minus recruiters...), 'Love the Steelers?' (Join other Steelers fans...), 'Professors Travel Free' (Professors, Show Your Students The World On A Faculty-led Tour...), and 'Ironman Lake Placid 2013' (Sign up and receive free registration...).

You can't talk about networking without thinking about economics anymore...

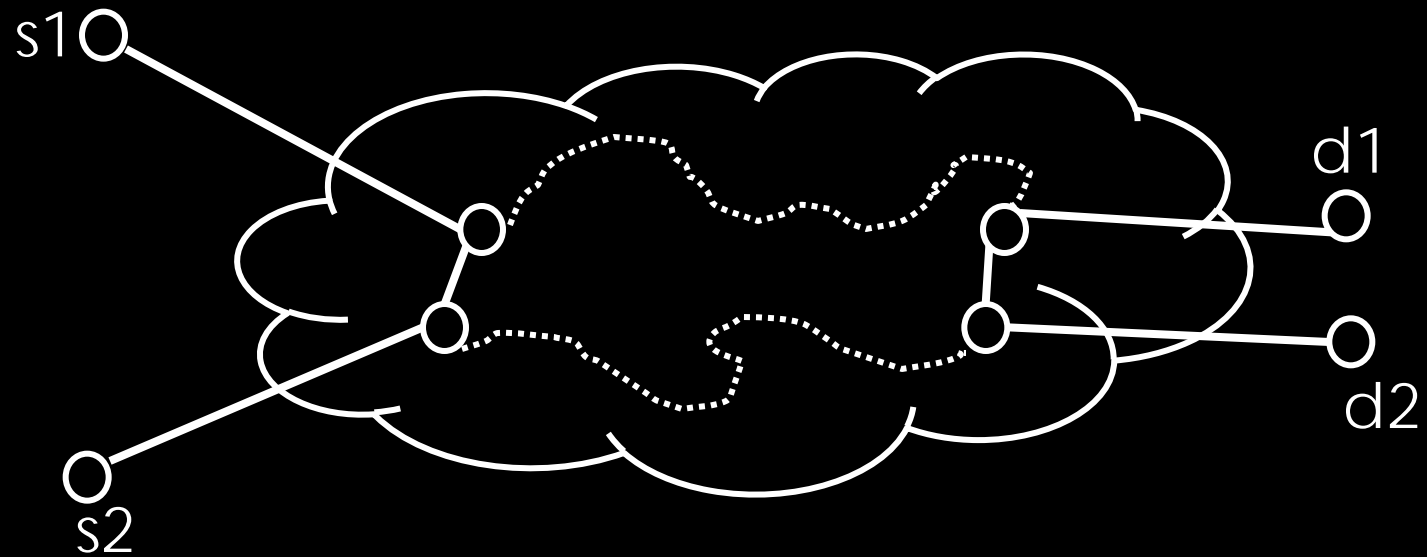


Invented at Caltech
(sorry!)

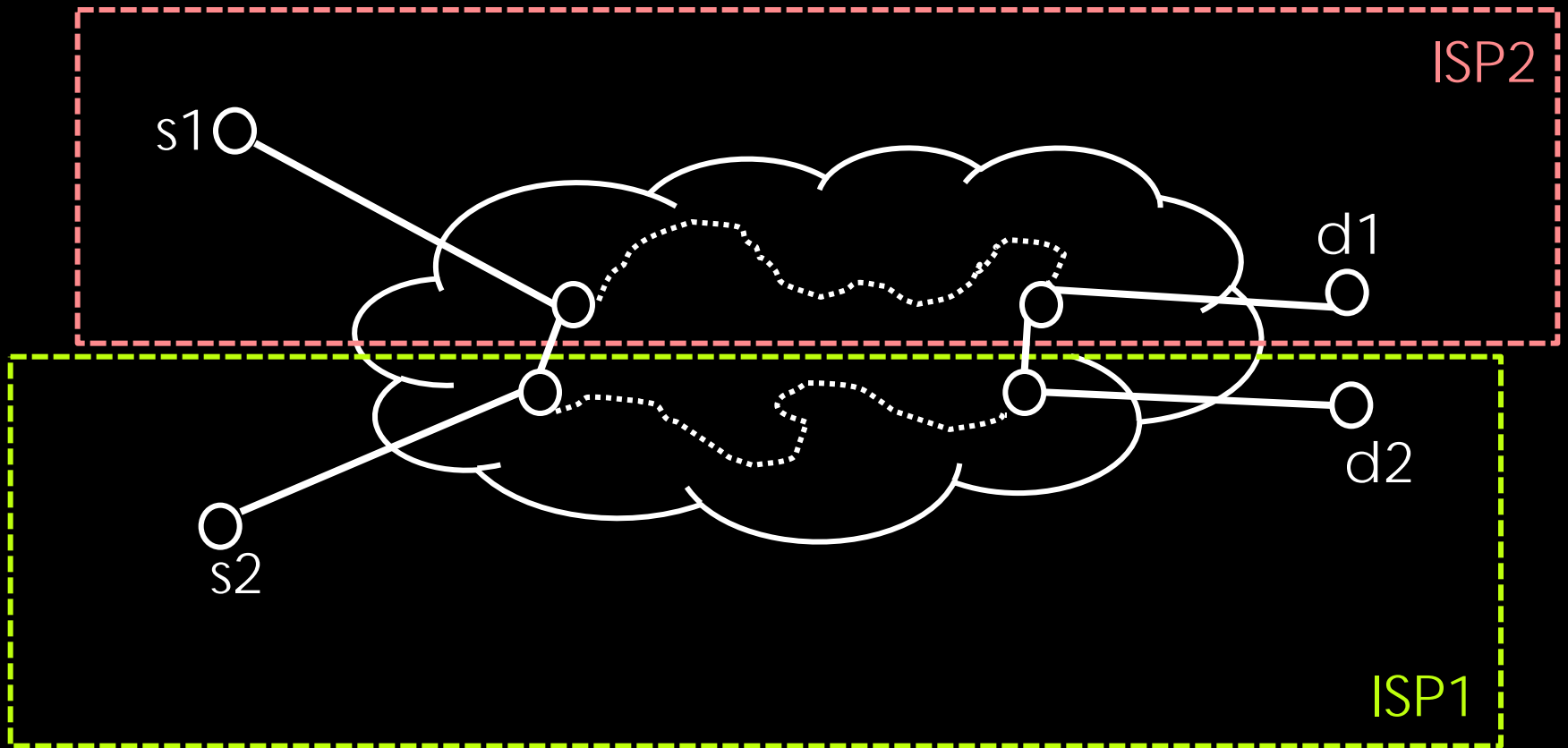


NET NEUTRALITY
ALL BITS ARE CREATED EQUAL
It's not just a good idea,
It ought to be the LAW.

How do ISPs make routing decisions?



How do ISPs make routing decisions?





Many startups exploit properties such as network formation, heavy-tails, auctions, centrality, virality, etc. that we will study in this class.

This course is about what the networks we
interact with “look” like

... how they “work”

... and how to exploit them!

COURSE OUTLINE

Understanding network structure

- Universal properties of networks
- What causes these properties?

Exploiting network structure

- How search works
- Virality & epidemics

Economics and networks

- Routing & ISP games
- Ad auctions

KEY GOAL:

Mix “theory” and “practice” throughout

This is a unique course

... that comes with pluses and minuses

- + modern material
- + can't take a course like this anywhere else
- + topics chosen to match research & industry focus
- + you have input into forming the course

- new material is being developed on the fly each year
- assignments may have bugs
- textbook only covers pieces of what I'll teach

WARNING:

This course...

- covers challenging material
- has time consuming homeworks & mini-projects
- requires you to both program & prove theorems

Administrivia

The staff

Professor

→ Adam Wierman

TAs

→ John Pang (lead)

→ Aditya Karan

→ Taokun (Xander) Zheng

→ Donsuk (Don) Lee

→ Yujie Tang

→ Xiaoqi Ren

The lectures

Mon & Wed & (some) Fri 9-10:30am
...Friday's will serve as "pre-make-up"
lectures for classes I will miss due to travel.

I know that this is early, but I really hope that you attend lectures anyway. The class works much better if everyone attends and interacts!

The course webpage

Course details

The course meets MW 9-10:25am in Annenberg 105.

We will often have "makeup" and "bonus" lectures on Friday 9-10:25am in Annenberg 105.

Note: The first lecture is Jan 6. There will not be a lecture on Jan 4.

The course will be managed using Piazza. All communication will happen and all materials will be posted through <http://piazza.com/caltech/winter2016/cseecms144>. Email Adam if you have problems enrolling yourself at the site.

Course Description

Social networks, the web, and the internet are an essential parts of our lives and we all depend on them every day, but do you really know what makes them work? This course studies the "big" ideas behind our networked lives. Things like, what do networks actually look like (and why do they all look so similar)? How do search engines work? Why do memes spread the way they do? How does web advertising work? For all these questions and more, the course will provide a mixture of both mathematical analysis and hands-on labs. *This course can be combined with CS/EE 145 and CS 142 or CS/EE 143 to satisfy the project requirement for CS undergraduate degree, but CS/EE 143 and CS 141a are not required prerequisites. The course assumes students are comfortable with graph theory, probability, and basic programming.*

Instructor

[Adam Wierman](mailto:adamw@caltech.edu), adamw@caltech.edu

The course webpage

PIAZZA

CMS/CS/EE 144 ▾ Q & A **Resources** Statistics Manage Class

Adam Wier

California Institute of Technology (Caltech) - Winter 2016

CMS/CS/EE 144: Networks: Structure & Economics

+ Add Syllabus

Course Information Staff **Resources**

Description

Edit

Social networks, the web, and the internet are an essential parts of our lives and we all depend on them every day, but do you really know what makes them work? This course studies the "big" ideas behind our networked lives. Things like, what do networks actually look like (and why do they all look seemingly look so similar)? How do search engines work? Why do memes spread the way they do? How does computational advertising work? For all these questions and more, the course will provide a mixture of both mathematical analysis and hands-on projects.

This course can be combined with CS/EE 145 and CS 142 or CS/EE 143 to satisfy the project requirement for CS undergraduate degree, but CS/EE 142 and CS 143 are not required prerequisites. The course assumes students are comfortable with graph theory, probability, and basic programming.

General Information

Edit

Course Details

We will meet weekly on some subset of MWF 9-10:25am in Annenberg 105. Because of my travel schedule, we will meet 2-3 times some weeks and 0-1

Announcements

+ Add

Lecture schedule

Edit Delete

1/02/15 11:34 AM

One quirk of the class this term is that we will not be strictly keeping to Monday/Wednesday lectures each week. There are a few points where I have travel constraints that will cause me to miss lectures, and so we are going to build up a buffer of extra lectures during the first half of the term (by adding a few Fridays) so that we still have the same number of lectures (and cover the same amount of content) as I have in previous years.

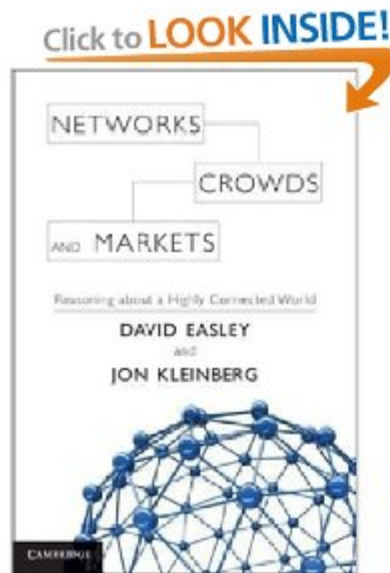
To help you keep track of which days lectures are happening, I will keep this post up-to-date with the lecture plan for the term. So, I suggest you pin this note and check back often.

Tentative lecture plan

Introduction to the class

"there's an app for that"

The textbook



Networks, Crowds, and Markets: Reasoning About a Highly Connected World [Hardcover]

[David Easley](#) (Author), [Jon Kleinberg](#) (Author)

★★★★★ (2 customer reviews)

Like (3)

List Price: ~~\$50.00~~

Price: **\$40.40** & this item ships for **FREE with Super Saver Shipping.** [Details](#)

You Save: **\$9.60 (19%)**

[Special Offers Available](#)

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered Tuesday, January 4? Order it in the next **26 hours and 6 minutes**, and choose **One-Day Shipping** at checkout. [Details](#)

29 new from \$40.40 **16 used** from \$34.93

- > You can get a pdf copy if you don't want to buy it
- > I will also post pdfs of my notes on piazza

Useful background

Basic Probability

Basic Graph Theory

Basic Game Theory (not required)

Programming experience (python)



We will have "bonus" lectures to refresh/teach these.

Probability refresher is THIS FRIDAY

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

** I reserve the right to add a final and adjust percentages accordingly

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

Homeworks will mix theory and practice.

They will hopefully be challenging and fun.

→ The first one goes out Wed and is due next Wed.

If you can do it in 6-9hrs, you have the necessary background.

→ Come to office hours.

...they will be in the evening (7-9) the 2 nights before the HW is due.

→ Mini-projects this year: rankmaniac, pandemaniac, clickmaniac

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

Solutions will be handed out the first lecture after homeworks are due. You will then have ~2 days to use the solutions to mark & correct your homework.

- We prefer electronic submissions, but if you turn HW in on paper be sure to make a photocopy of your homework before you turn it in!
- The TAs grade your homework independently of your “self-evaluation”, and then use their grading to verify your self-evaluation.

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

We will have unannounced in-class quizzes 5-10 times during the term. These will be easy, so I do not expect you to study, they will be meant only to verify your understanding of basic definitions in the class (and to reward those that attend lecture).

-- 50% will be guaranteed just by being present for the quiz!

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%



Project proposal – 10%

Project plan – 10%

Blog post – 4%

We won't be doing a project, we'll just be preparing for the project in the third term.

- Proposal: Describe three ideas in 1-2 pages in a public post on piazza so that you can get feedback from the class. You do this whether you're taking 145 or not.
- Plan: Describe your detailed plan for 145 in 5-10 pages. If you're not taking 145 you design a "course module" instead.

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%



Project proposal – 10%

Project plan – 10%

Blog post – 4%

I'm expecting lots of great ideas!

Over the past few years we've had, among other things,

-- a patent filed

-- a startup incorporated

-- a paper accepted to a conference

-- an idea "crushed" by facebook

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

Everyone is required to post one short blog post during the term.
See Piazza for details...

→ These are useful in finding ideas for projects!

Your grade

Homeworks & mini-projects – 60%

Homework self-evaluations – 8%

Quizzes – 8%

Project proposal – 10%

Project plan – 10%

Blog post – 4%

Any questions?

Course warmup...

The evolution and structure of the web



Does anyone know when the web was invented?

→ 1989

...but the ideas that led to the
web are from much earlier

The web is an implementation of an information network



A graph where nodes are pieces of information and edges link related pieces of information


What was new and exciting about the web?

The web is an implementation of an information network

This is what was
new & exciting



Information networks have
been used & studied for ages.
e.g. citation networks,
cross-refs in encyclopedias



1945



Vannevar Bush

As We May Think - The Atlantic (July 1945) - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.theatlantic.com/doc/194507/bush

Internet Archive Wayback Machine - W... As We May Think - The Atlantic (...)

by *Vannevar Bush*

As We May Think

Director of Scientific Research and Development, Dr. Bush coordinated the activities of some six thousand scientists in the application of science to warfare. In this article he holds up an incentive for scientists when the war ends: that men of science should then turn to the task of making more accessible our bewildering store of knowledge. His inventions have extended man's physical powers and sharpened his mind. Trip hammers that multiply the fists, microscopes that reveal the hidden, engines of destruction and detection are new results, but not the end of the road. Now, says Dr. Bush, instruments are at hand which, if properly used, will give man access to and command over the inherited knowledge of the ages. He argues that peaceful instruments should be the first objective of our scientists as war work ends. Like Emerson's famous address of 1837 on "The American Scholar," Dr. Bush calls for a new relationship between thinking man and the machine.

—THE EDITOR

ARTICLE TOOLS
powered by:
The Atlantic
E-MAIL ARTICLE
PRINTED NORMAL

This has not been a scientist's war; it has been a war in which all have had a part. The scientists, burying their old professional competition in the demand of a common cause, have shared greatly and learned much. It has been exhilarating to work in effective partnership. Now, for many, this appears to be approaching an end. What are the scientists to do next?

For the biologists, and particularly for the medical scientists, there can be little indecision, for their war has hardly required them to leave the old paths. Many indeed have been able to carry on their war research in their familiar peacetime laboratories. Their objectives remain much the same.

It is the physicists who have been thrown most violently off stride, who have left academic pursuits for the making of strange destructive gadgets, who have had to devise new methods

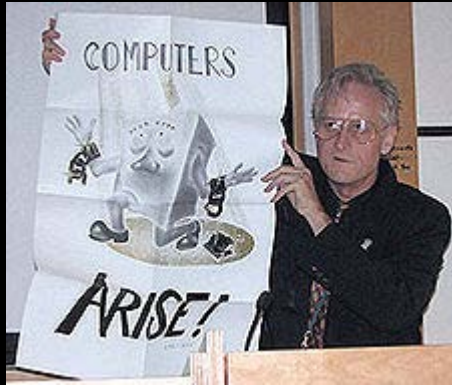
Done

1945



- 1) When we store information in books,
it is highly linear
 - 2) When we think, we use a
semantic network
- Technology will allow us to digitize
knowledge and mimic the
semantic network connections.

1965

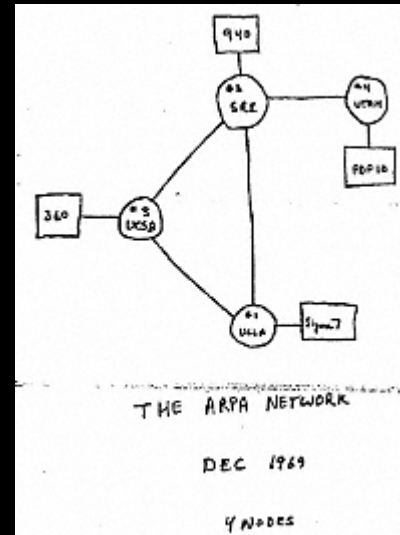


Ted Nelson coins the
term "hypertext"

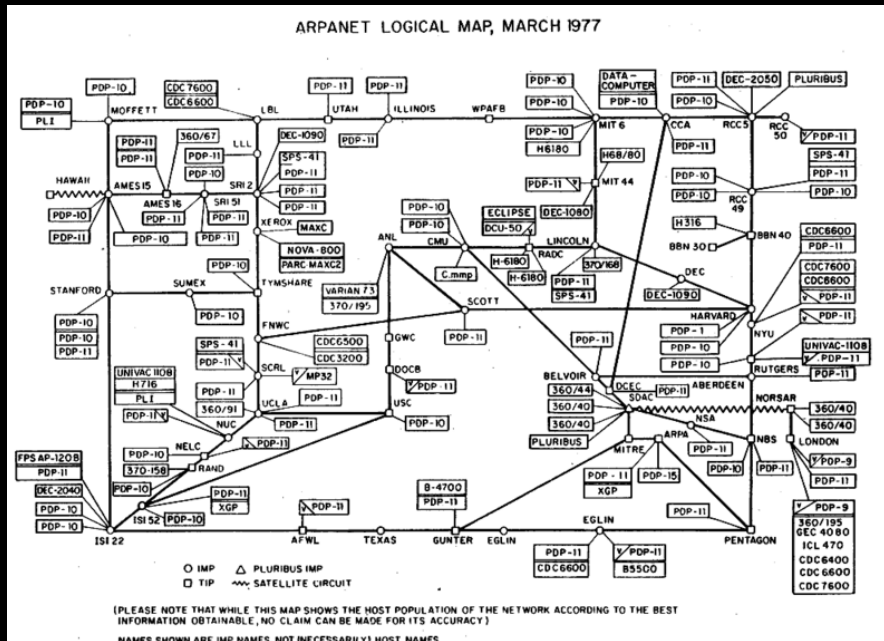
↖
A computer assisted form of authoring
meant to replace linear structure of text
with a network.

1969

1969→



←1977



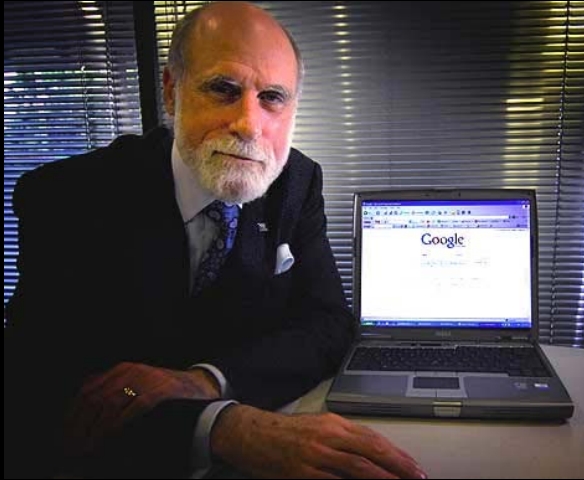
ARPANET begins...with a deployment at UCLA, Stanford, UCSB, and Utah

1971



Ray Tomlinson creates first email program

1974



TCP was defined by Vint Cerf & Bob Kahn

1978

TCP splits to become TCP & IP

1984



Paul Mockapetris introduces DNS

1989 – The web emerges



The original proposal of the WWW, HTMLized - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.w3.org/History/1989/proposal.html

Information Management: A Proposal

Tim Berners-Lee, CERN
March 1989, May 1990

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Overview

Many of the discussions of the future at CERN and the LHC era end with the question - "Yes, but how will we ever keep track of such a large project?" This proposal provides an answer to such questions. Firstly, it discusses the problem of information access at CERN. Then, it introduces the idea of linked information systems, and compares them with less flexible ways of finding information.

It then summarises my short experience with non-linear text systems known as "hypertext", describes what CERN needs from such a system, and what industry may provide. Finally, it suggests steps we should take to involve ourselves with hypertext now, so that individually and collectively we may understand what we are creating.

X Find: loader Next Previous Highlight All Match case

Done

Tim Berners-Lee writes "Information Management: A proposal" at CERN

The goal was to improve project management by building a network interface for hypertext.

1989

The original proposal of the WWW, HTMLized - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.w3.org/History/1989/proposal.html

The original proposal of the WWW, H...

What will the system look like?

Let us see what components a hypertext system at CERN must have. The only way in which sufficient flexibility can be incorporated is to separate the information storage software from the information display software, with a well defined interface between them. Given the requirement for network access, it is natural to let this clean interface coincide with the physical division between the user and the remote database machine.

This division also is important in order to allow the heterogeneity which is required at CERN (and would be a boon for the world in general).

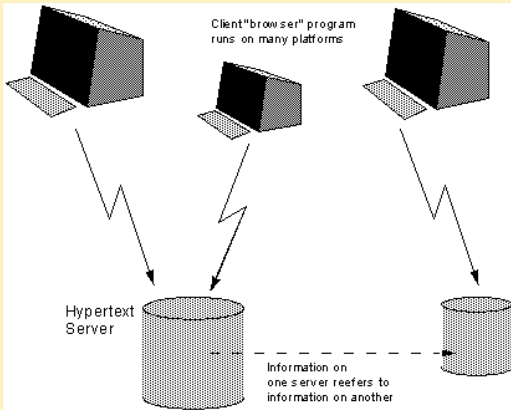
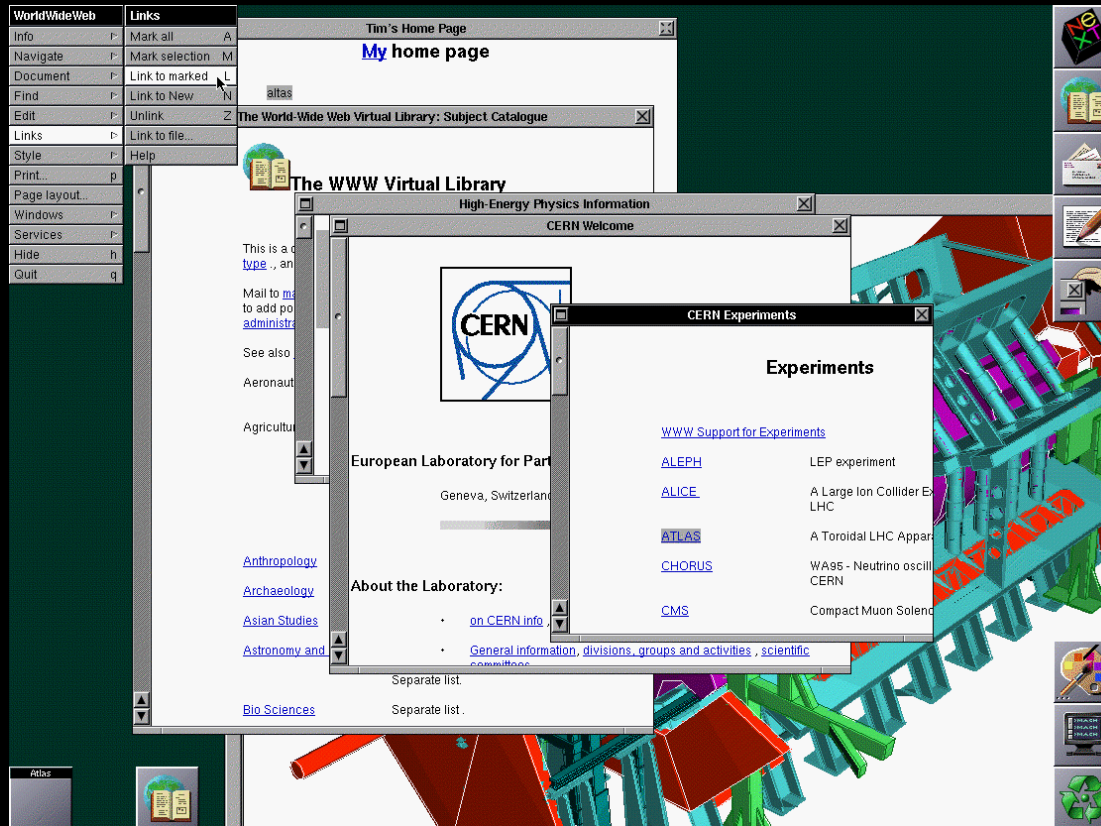


Fig 2. A client/server model for a distributed hypertext system.

Therefore, an important phase in the design of the system is to define this interface. After that, the development of various forms of display program and of database server can proceed in parallel. This will have been done well if many different information sources, past, present and future, can be mapped onto the definition, and if many different human interface programs can be written over the years to take advantage of new technology and standards.

Done

1990



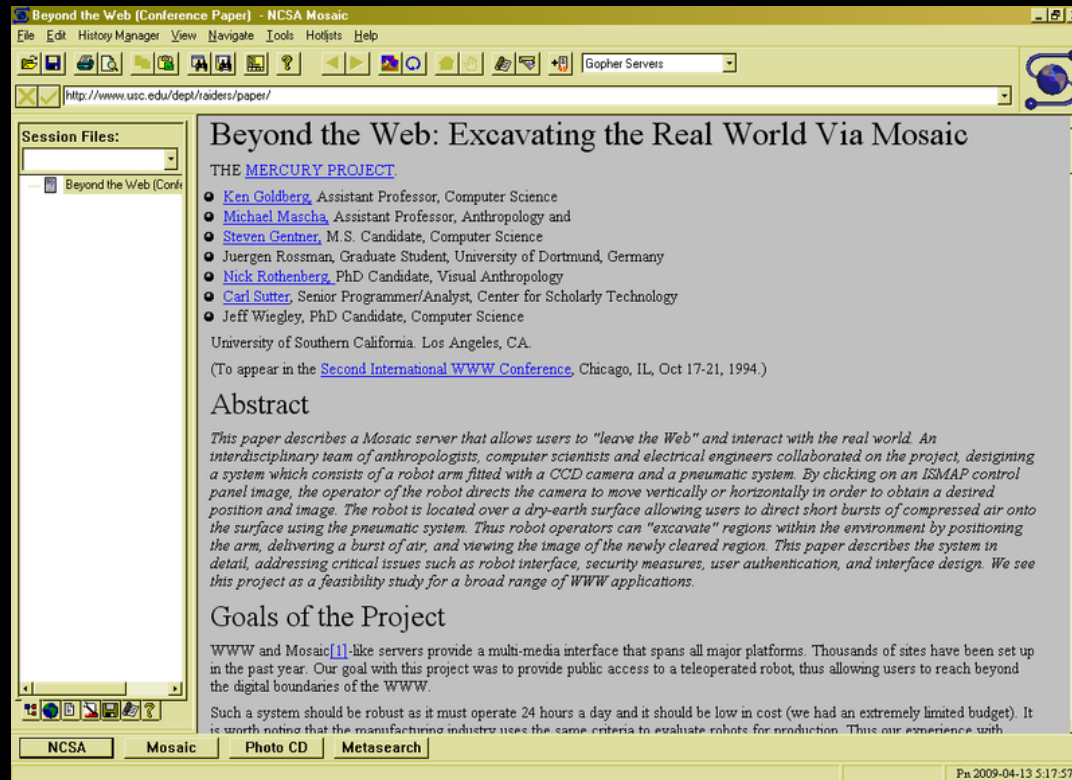
First browser developed at CERN

1991



First paper appears on the project at
Hypertext conference
→ Only accepted as a poster!

1993



Mosaic became the first graphical browser
(My first browser!)

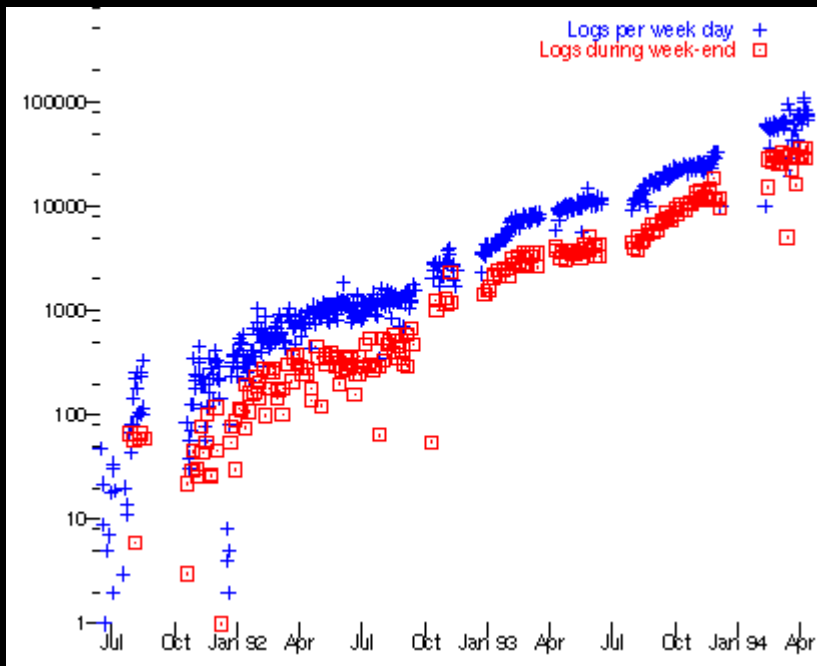
1993

CERN agrees to allow public use of
web protocol royalty-free!

1994

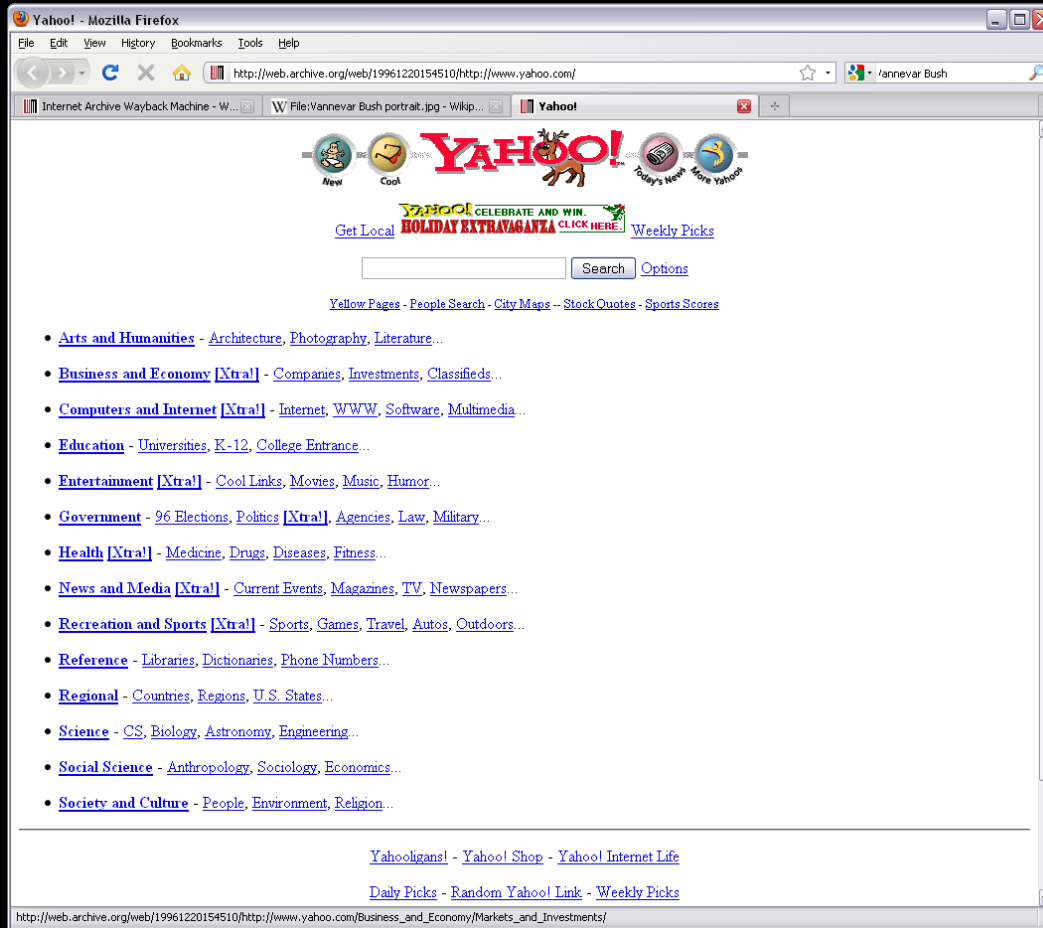
→ Mosaic goes commercial
(later becomes Netscape)

→ Traditional dialups (AOL, CompuServe, Prodigy) begin to sell Internet access.



Load on 1st
web server

1994

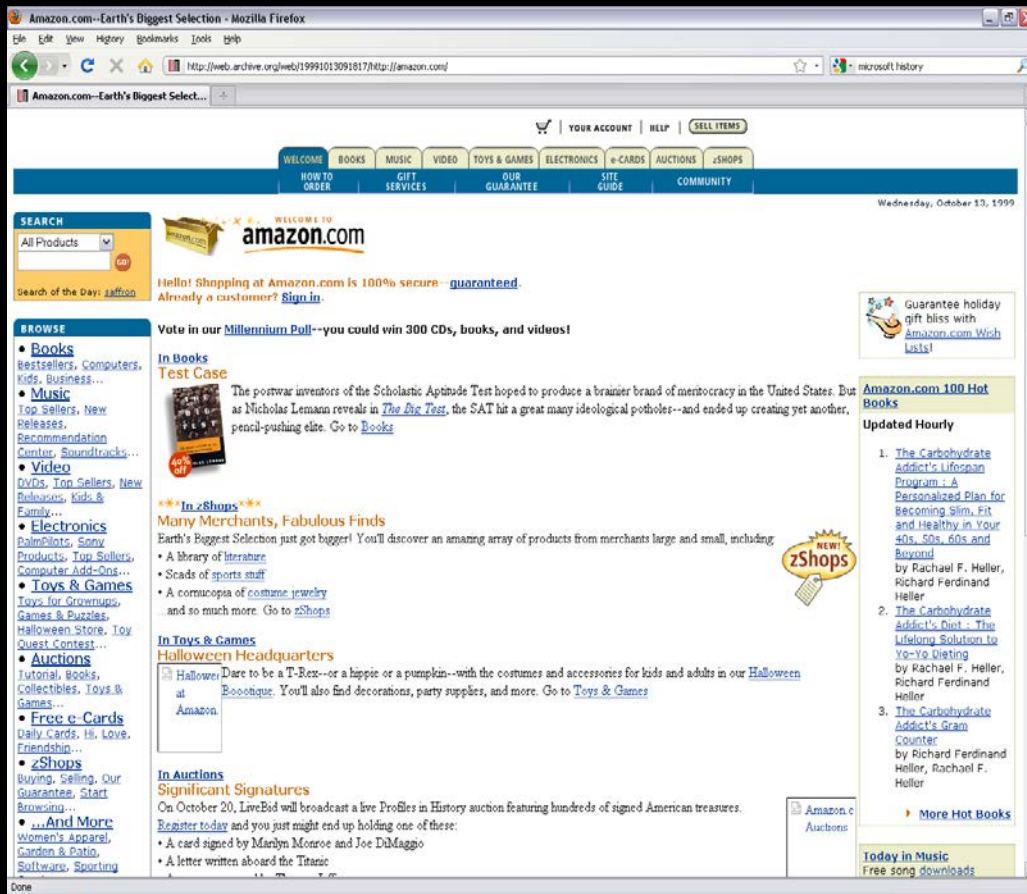


Yahoo
circa
1996

“Jerry’s Guide to the world wide web”
started ... it eventually became Yahoo

1995+

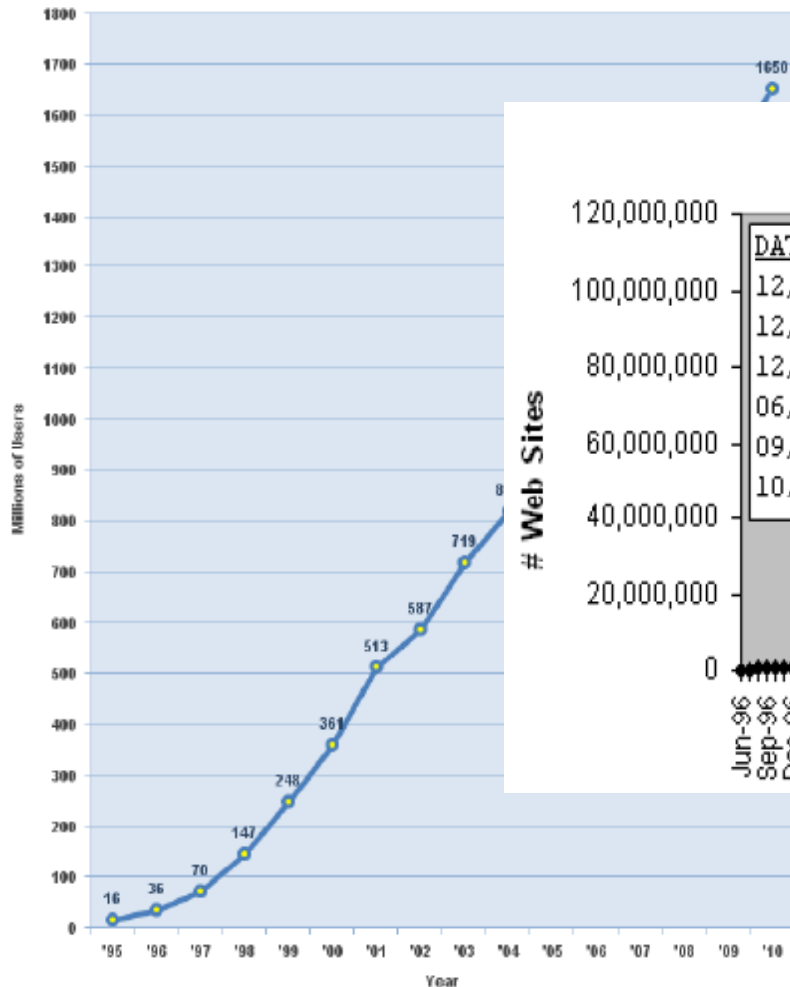
Amazon arrives and the commercialization of the web begins



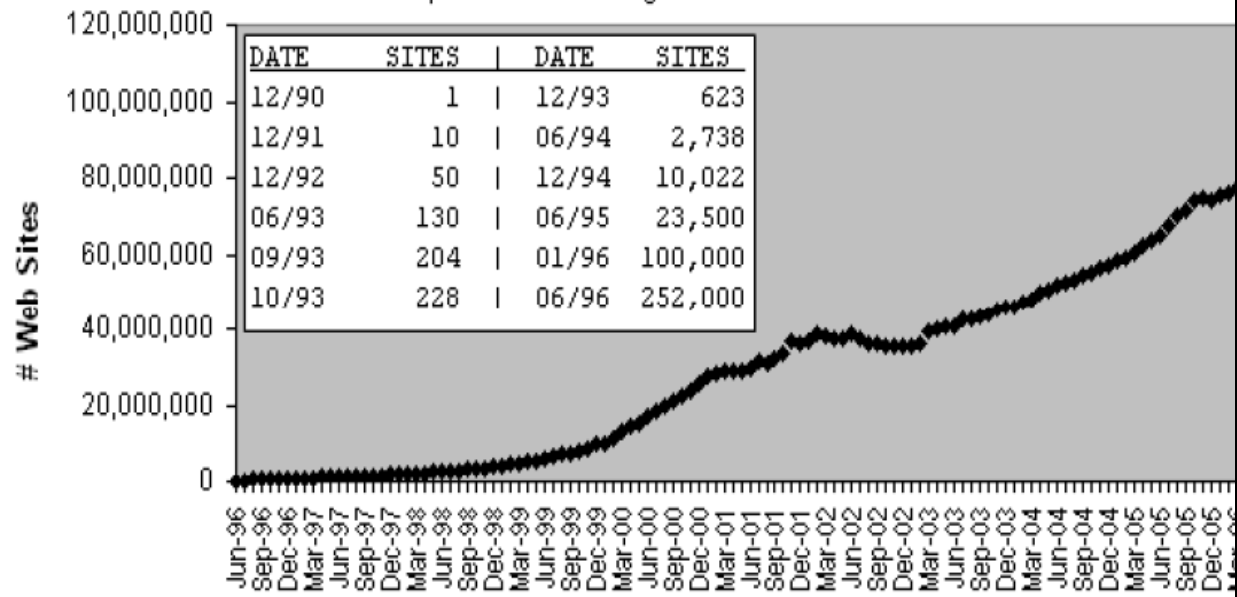
Amazon
circa
1999

The growth of the web

Internet Users in the World
Growth 1995 - 2010



Hobbes' Internet Timeline Copyright ©2006 Robert H Zakon
<http://www.zakon.org/robert/internet/timeline/>



The growth of the web

3 data points about the google index

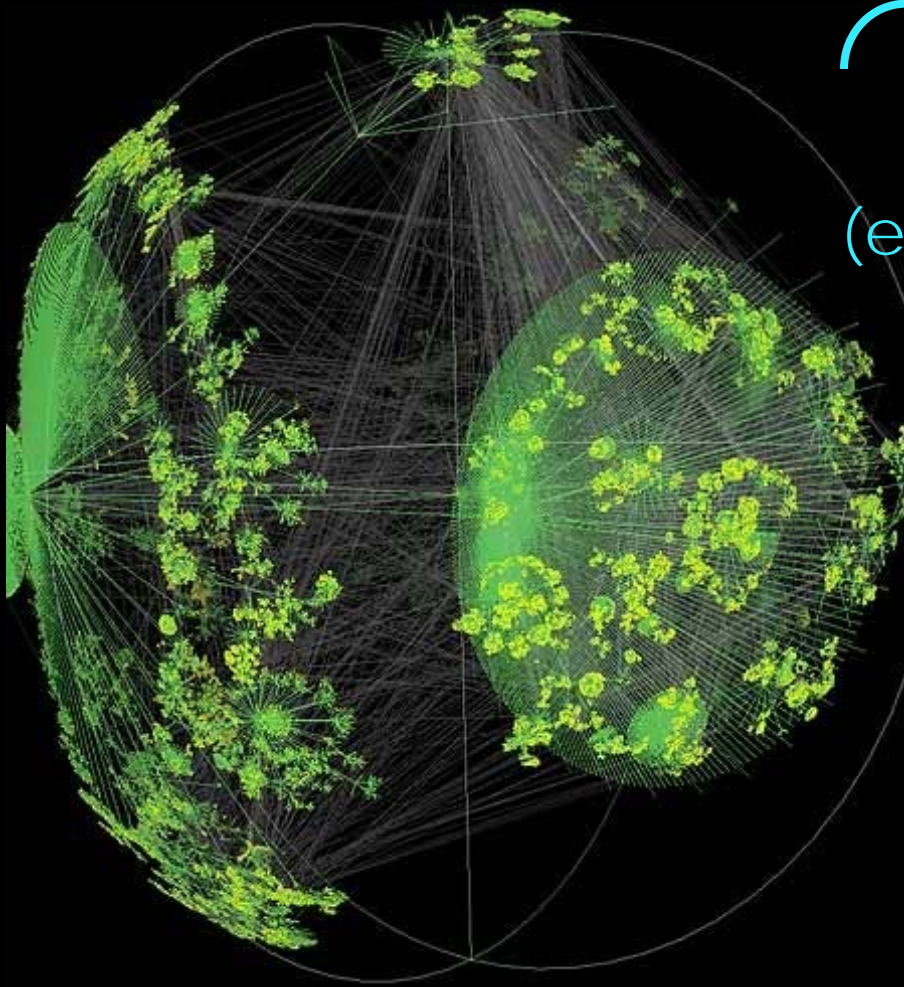
1998 → 28 million pages

2000 → 1 billion pages

2008 → 1 trillion pages

2014 → so big that it doesn't matter any more...

Our goal: What “structure” does the web graph have?



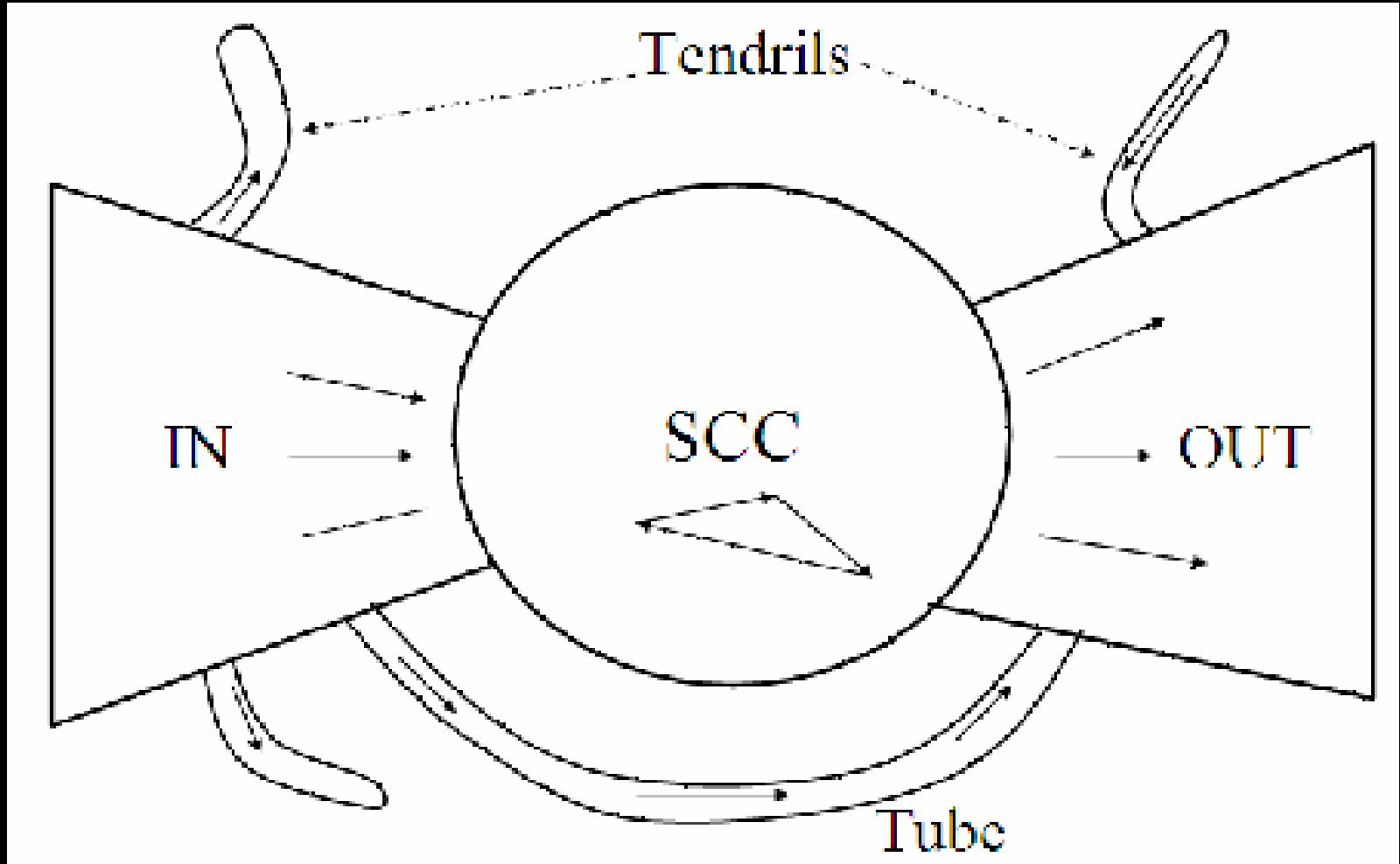
Nodes = pages
Edges = links
(edges are directed)

Our goal: What “structure” does the web graph have?

What properties should we look at?

Our goal: What “structure” does the web graph have?

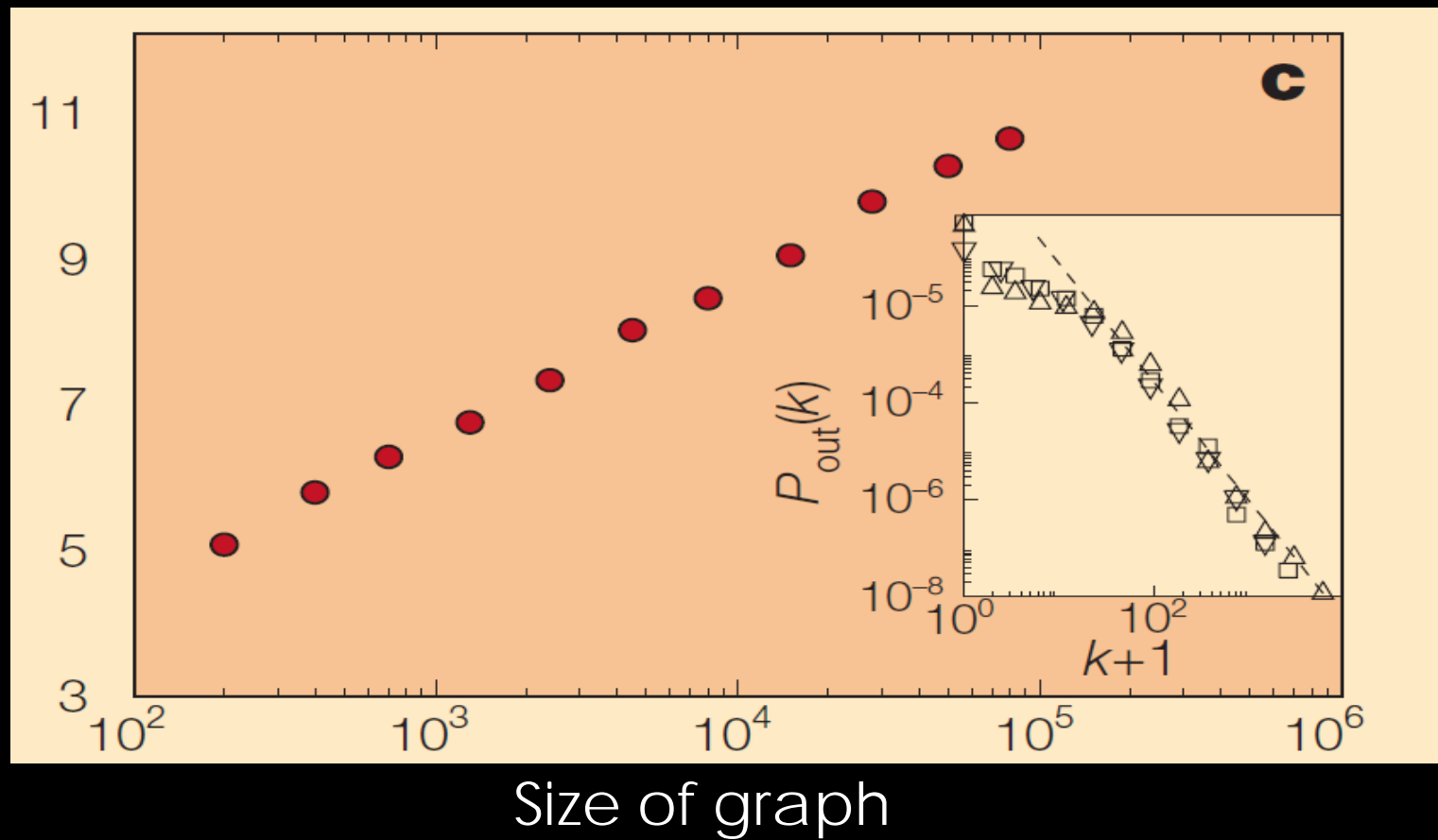
- 1) Connectivity?
- 2) Diameter?
- 3) Degree?
- 4) Clustering?



Our goal: What “structure” does the web graph have?

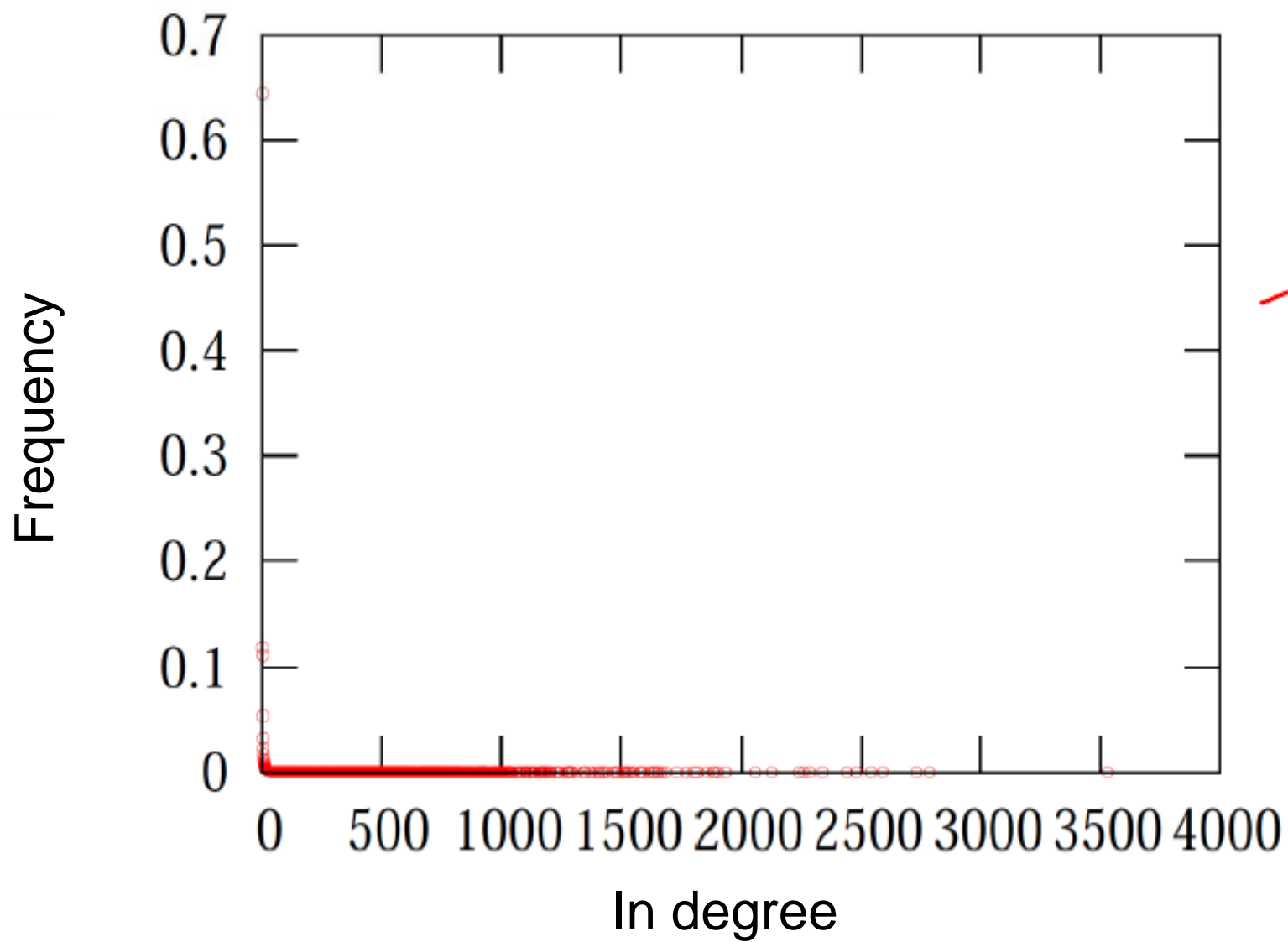
- 1) Connectivity? There is a giant connected component
- 2) Diameter?
- 3) Degree?
- 4) Clustering?

Average distance

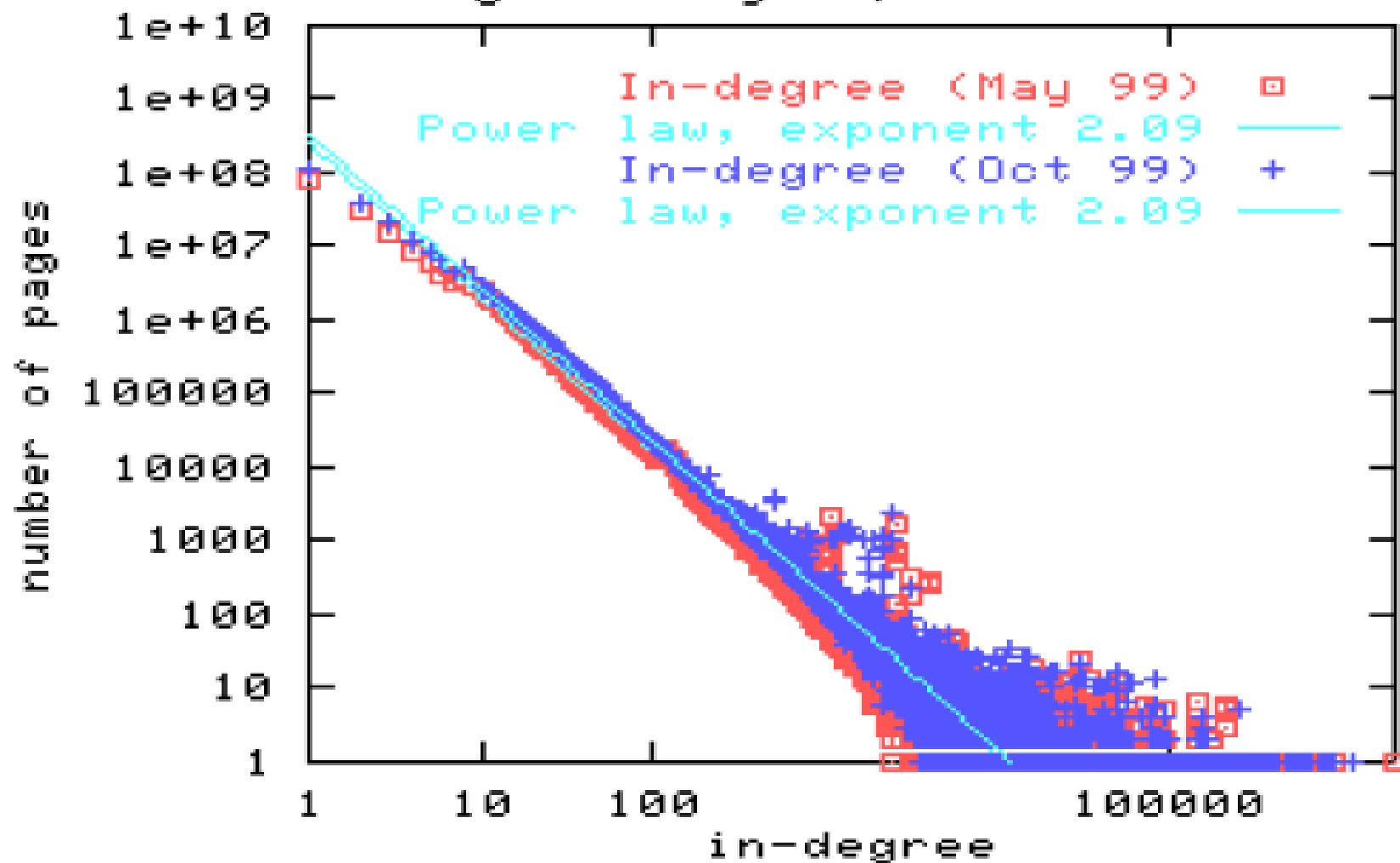


Our goal: What “structure” does the web graph have?

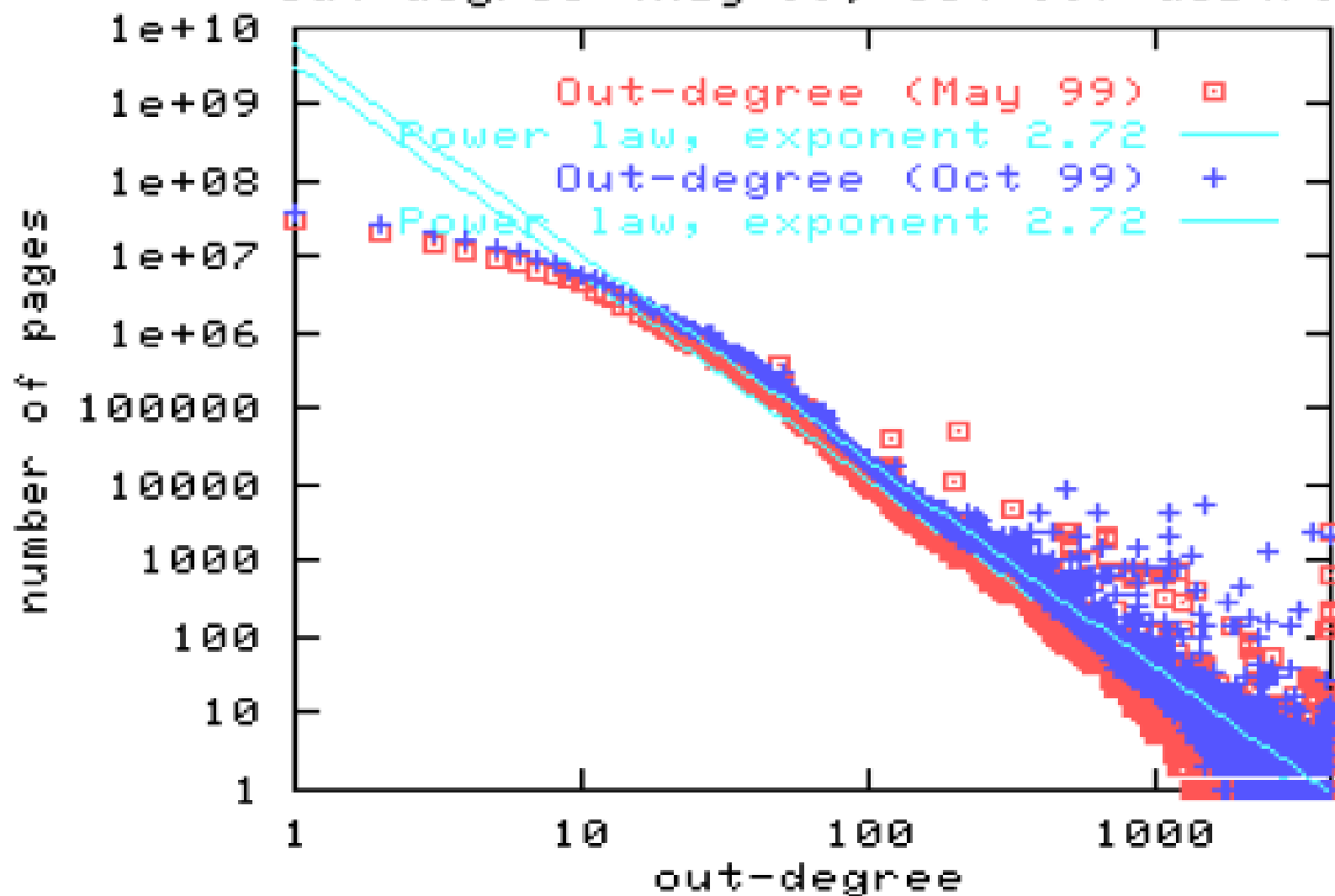
- 1) Connectivity? There is a giant connected component
- 2) Diameter? Small diameter
- 3) Degree?
- 4) Clustering?



In-degree (May 99, Oct 99) distr.



Out-degree (May 99, Oct 99) distr.

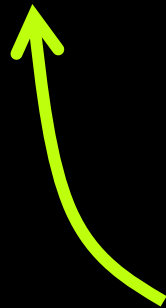


Our goal: What “structure” does the web graph have?

- 1) Connectivity? There is a giant connected component
- 2) Diameter? Small diameter
- 3) Degree? Heavy-tailed degree distribution
- 4) Clustering?

Our goal: What “structure” does the web graph have?

- 1) Connectivity? There is a giant connected component
- 2) Diameter? Small diameter
- 3) Degree? Heavy-tailed degree distribution
- 4) Clustering? Highly clustered



Next we'll see that these are “universal”