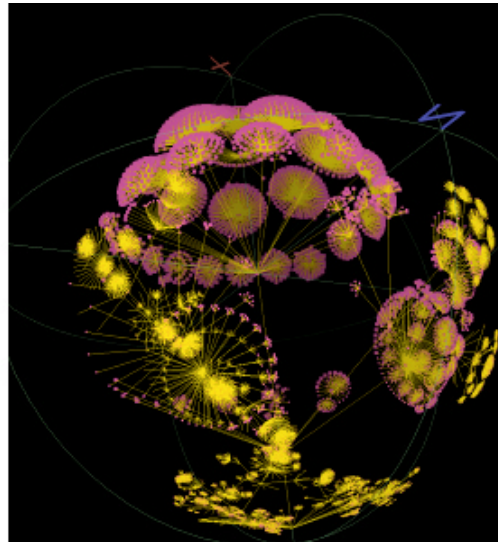


The power of maps in e-research



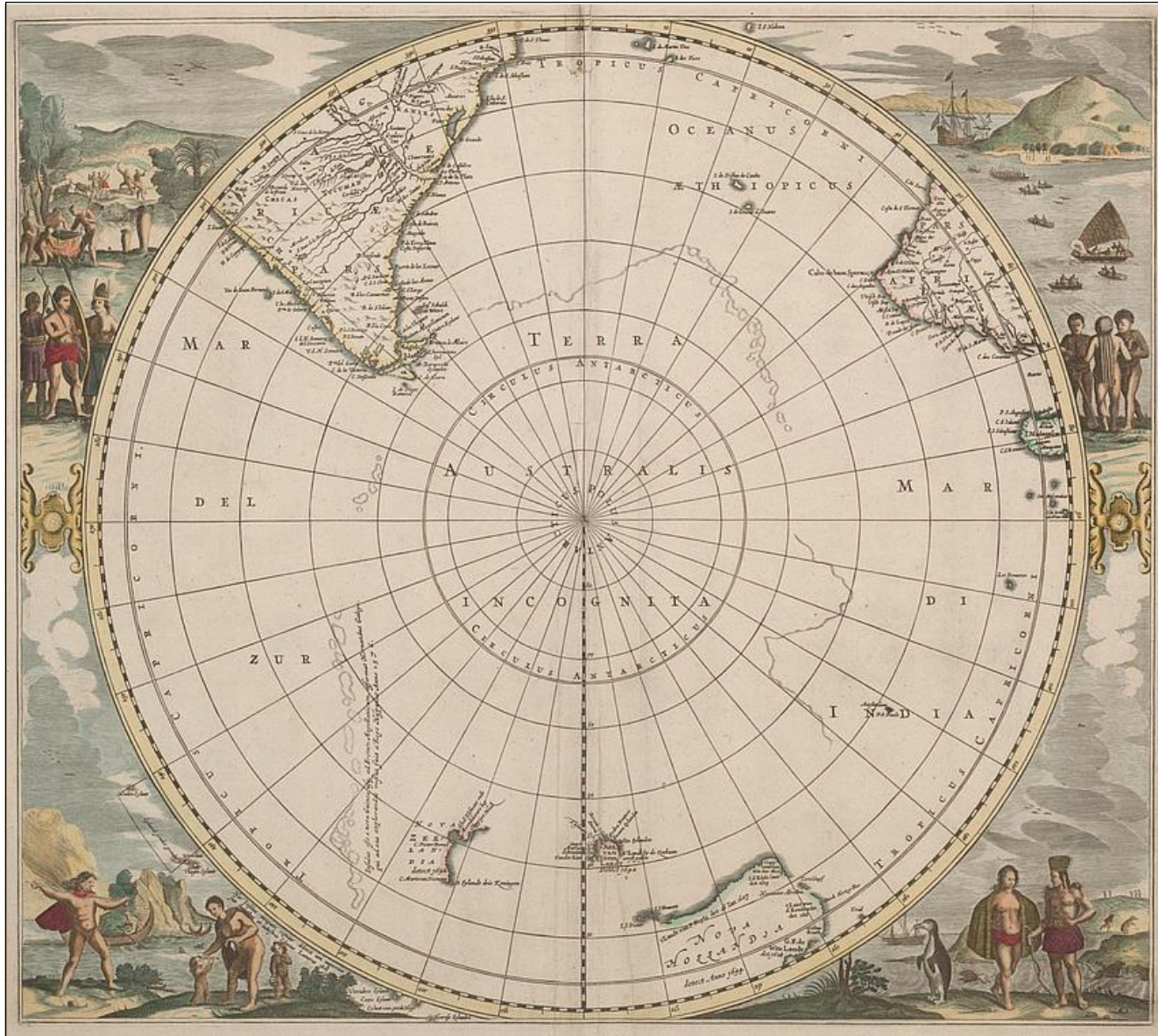
Martin Dodge

Centre for Advanced Spatial Analysis, University College London

e-Social Science: Enabling Technologies and New Research, 12th November 2004
The Australian National University, Canberra

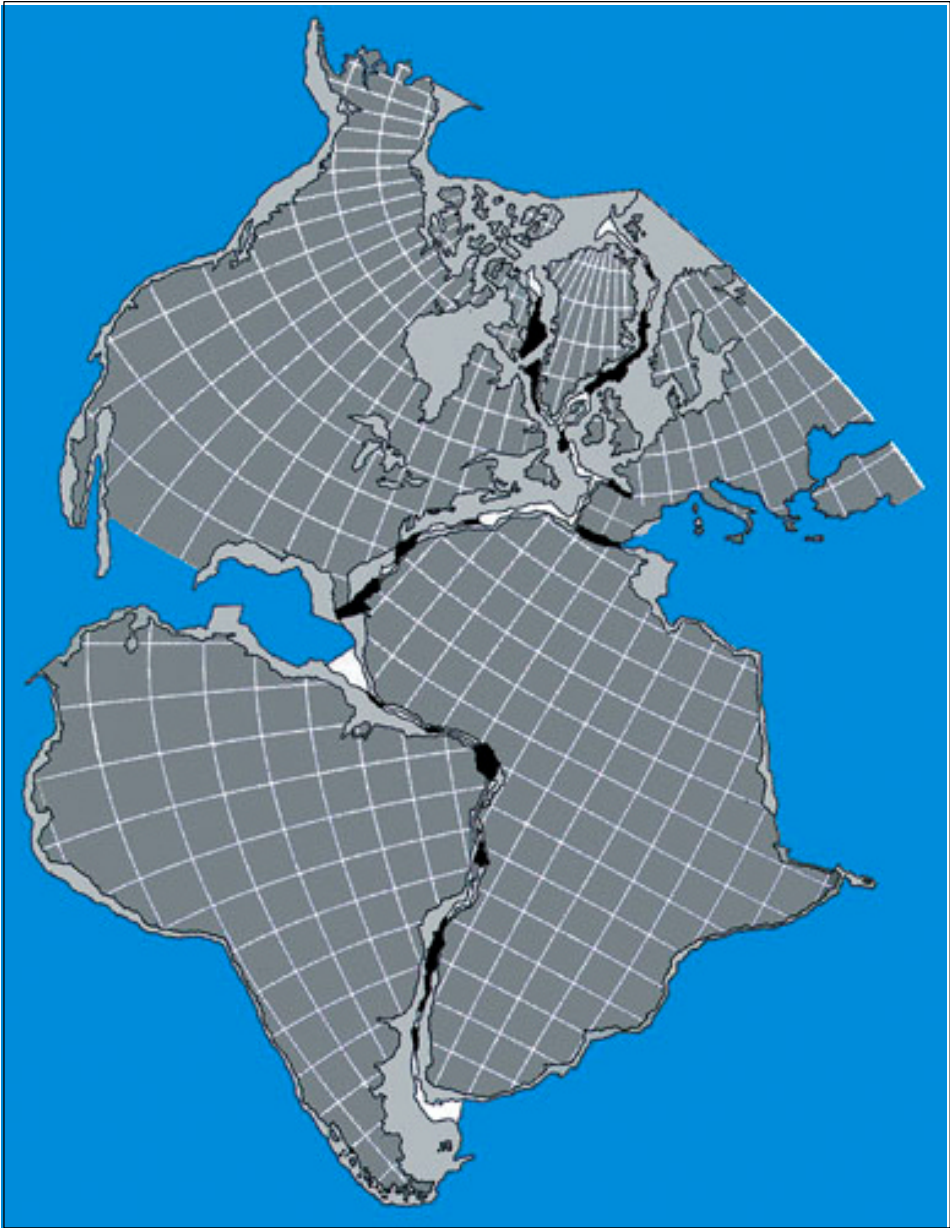
the map is a help provided to the imagination through the eyes

Henri Abraham Chatelain, Atlas Historique (1705)



Power of maps (in e-research)

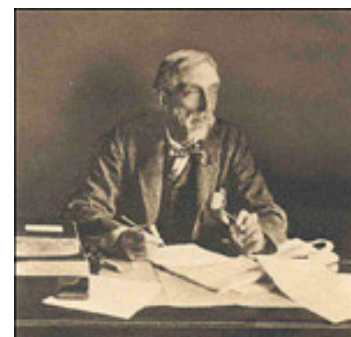
1. seeing patterns (eureka moments)
2. revealing the process (and proving theory to others)
3. classifying space and people (thus enabling social and spatial targeting of some response)



1898-99 Map



▲ Reduce view area



2001, All Rights Reserved: ULL

The Streets are coloured according to the general condition of the inhabitants, as under:-

- | | |
|--|--|
|  Lowest class. Vicious, semi-criminal. |  Fairly comfortable. Good ordinary earnings. |
|  Very poor, casual. Chronic want. |  Middle-class. Well-to-do. |
|  Poor. 18s. to 21s. a week for a moderate family. |  Upper-middle and Upper classes. Wealthy. |
|  Mixed. Some comfortable, others poor. | |

A combination of colours— as dark blue and black, or pink and red— indicates that the street contains a fair proportion of each of the classes represented by the respective colours.

1. Network geometries
2. Information topologies
3. Social structures

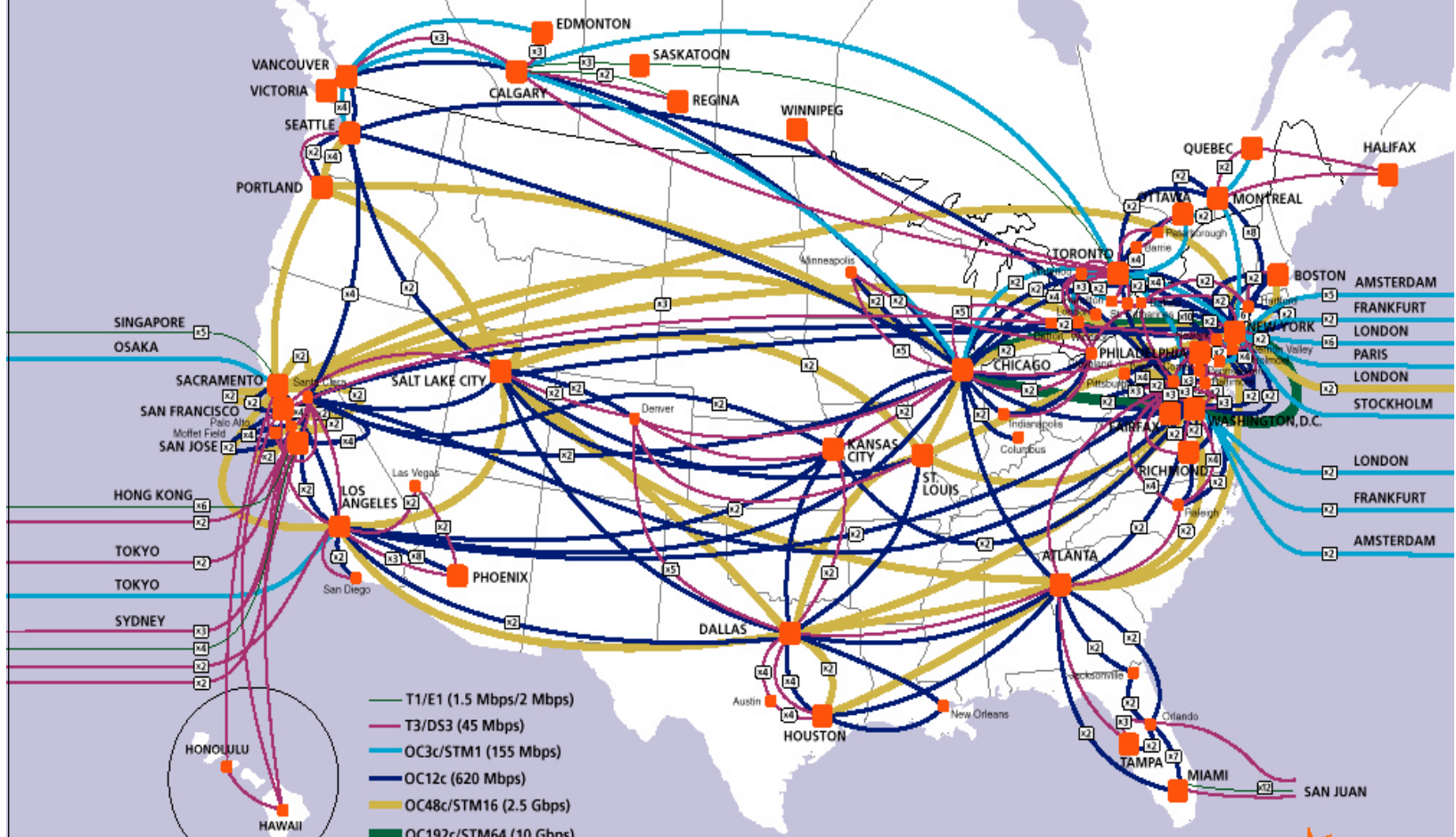
1. Network Geometries of the Internet infrastructures



Showing you where the wires go



UUNET's North America Internet network




- T1/E1 (1.5 Mbps/2 Mbps)
- T3/DS3 (45 Mbps)
- OC3c/STM1 (155 Mbps)
- OC12c (620 Mbps)
- OC48c/STM16 (2.5 Gbps)
- OC192c/STM64 (10 Gbps)
- Single Hub City
- Multiple Hubs City

For more information visit www.uu.net

NB: UUNET also has infrastructure within individual countries, which is not shown on this map.
Jan. 2000



advertisement

	Save Up To 70%		\$250,000		\$500,000		CLICK NOW FOR THE BEST RATES! <i>ReliaQuote</i> A Better Way to Buy Life Insurance
	10 Year Level Term Life Insurance Male/Female Monthly Premiums No Nicotine <small>LEGAL INFO</small>	AGE	MALE	FEMALE	MALE	FEMALE	
	35	\$10.22	\$9.14	\$16.10	\$13.92		
	45	\$18.49	\$15.44	\$32.63	\$26.54		

Sample rates underwritten by First Penn-Pacific Life Insurance Company.

TechNews.com Sign Up: [Free Daily Tech E-letter](#)

Technology Home

Washtech

Tech Policy

- Copyright
- Cybercrime
- E-Taxes
- FCC
- ICANN
- Security

Government IT

Markets

Columnists

Personal Tech

Special Reports

Jobs

washingtonpost.com

More News

advertisement

UPDATED DAILY!

Get Mortgage Rates from over

80

Local Lenders!

Loans Below **\$322,700**

Loans Above **\$322,700**

Dissertation Could Be Security Threat

Student's Maps Illustrate Concerns About Public Information

By *Laura Blumenfeld*

Washington Post Staff Writer

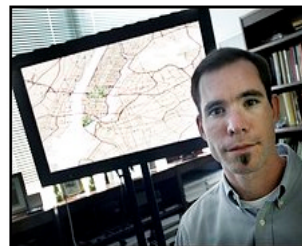
Tuesday, July 8, 2003; Page A01


Sean Gorman's professor called his dissertation "tedious and unimportant." Gorman didn't talk about it when he went on dates because "it was so boring they'd start staring up at the ceiling." But since the Sept. 11, 2001, attacks, Gorman's work has become so compelling that companies want to seize it, government officials want to suppress it, and al Qaeda operatives -- if they could get their hands on it -- would find a terrorist treasure map.

Tinkering on a laptop, wearing a rumpled T-shirt and a soul patch goatee, this George Mason University graduate student has mapped every business and industrial sector in the American economy, layering on top the fiber-optic network that connects them.

He can click on a bank in Manhattan and see who has communication lines running into it and where. He can zoom in on Baltimore and find the choke point for trucking warehouses. He can drill into a

advertisement

[enlarge photo](#) 

Sean Gorman's program can map critical infrastructure in cities. (Andrea Bruce Woodall -- The Washington Post)

the big unknowns - what is in those wires...

- there are no traffic maps
- no one really can tell you how much traffic is flowing across the Internet. no one knows how much, or from where to where
- limits of 'top-down' information
- power of internet to measure itself ('bottom-up' scanning)
- dynamic maps. real-time conditions
- traceroutes and weather maps

Report for www.ordsvy.gov.uk [62.25.96.204]

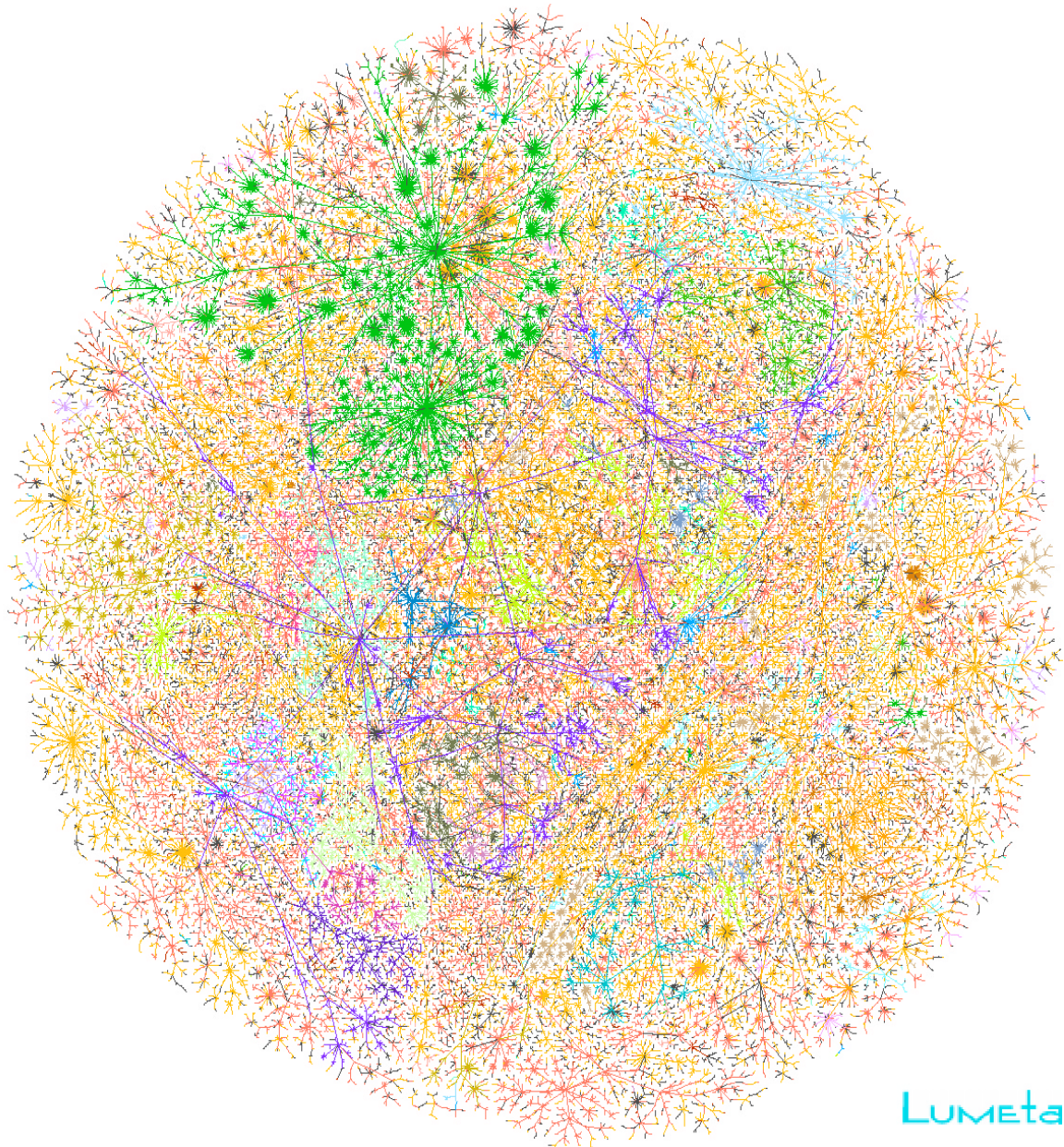
Analysis: www.ordsvy.gov.uk was found in 19 hops (TTL=235). It is a HTTP server (running Microsoft-IIS/5.0).

Hop	%Loss	IP Address	Node Name	Location	Tzone	ms	Graph	Network
0		203.94.129.69	webbms1.brd.n	-			0 313	United Energy Comm
1		203.43.193.22	gw-router.brd.r	-		0		Telstra Internet
2		203.94.129.25	-	(Australia)	*	0		United Energy Comm
3		203.94.128.41	fa0-1-2.que31-	-		0		Uecomm Australia
4		203.94.132.57	atm2-0-6.cla5f	-		0		United Energy Comm
5		203.208.148.2	-	Singapore	-02:00	0		SingTel Internet Exch
6		203.208.171.2	-	Singapore	-02:00	0		SingTel Internet Exch
7		203.208.154.4	-	Singapore	-02:00	156		SingTel Internet Exch
8		203.208.172.4	g0-0-0.plapx-a	Singapore	-02:00	156		SingTel Internet Exch
9		209.245.146.1	interconnect-ei	San Jose, CA, USA	-18:00	156		Level 3 Communicati
10		209.245.146.1	unknown.Level	-		156		Level 3 Communicati
11		64.159.2.33	ae0-52.mp2.S:	San Jose, CA, USA	-18:00	156		Level 3 Communicati
12		64.159.0.229	so-3-0-0.mp1.1	Washington, DC, U	-15:00	234		Level 3 Communicati
13		212.187.128.1	so-2-0-0.mp1.1	London, UK	-10:00	296		Transatlantic cable s
14		212.187.128.5	so-2-0-0.mp1.1	London, UK	-10:00	296		European MPLS core
15		212.187.131.8	gige10-0.ipcol	London, UK	-10:00	296		2nd London1 Gatewa
16		195.50.116.19	-	(United Kingdom)	-10:00	302		customer serial links
17		195.92.201.24	hnr-1.ewfd.asf	(United Kingdom)	-10:00	312		In case of problems,
18		62.25.96.194	-	(United Kingdom)	-10:00	312		Energis Squared Wa
19		62.25.96.204	www.ordsvy.gc	(United Kingdom)	-10:00	312		Energis Squared Wa



Legend

■ cw.net	6070
■ alter.net	3997
■ sprintlink.net	2479
■ att.net	2294
■ apnic.net	2219
■ ripe.net	2032
■ ans.net	1843
■ uu.net	1545
■ bbnplanet.net	1438
■ qwest.net	1243
■ telstra.net	1120
■ psi.net	1120
■ verio.net	1056
■ krmic.net	897
■ bellsouth.net	866
■ gblx.net	688
■ teleglobe.net	586
■ gjp.net	581
■ level3.net	536
■ pnap.net	514
■ digex.net	510
■ exodus.net	496
■ swbell.net	431
■ uswest.net	422
■ savvis.net	375
■ icix.net	374
■ kpnqwest.net	326
■ cerf.net	307
■ pbi.net	305
■ other ISPs	32871
■ not an ISP	
■ error	



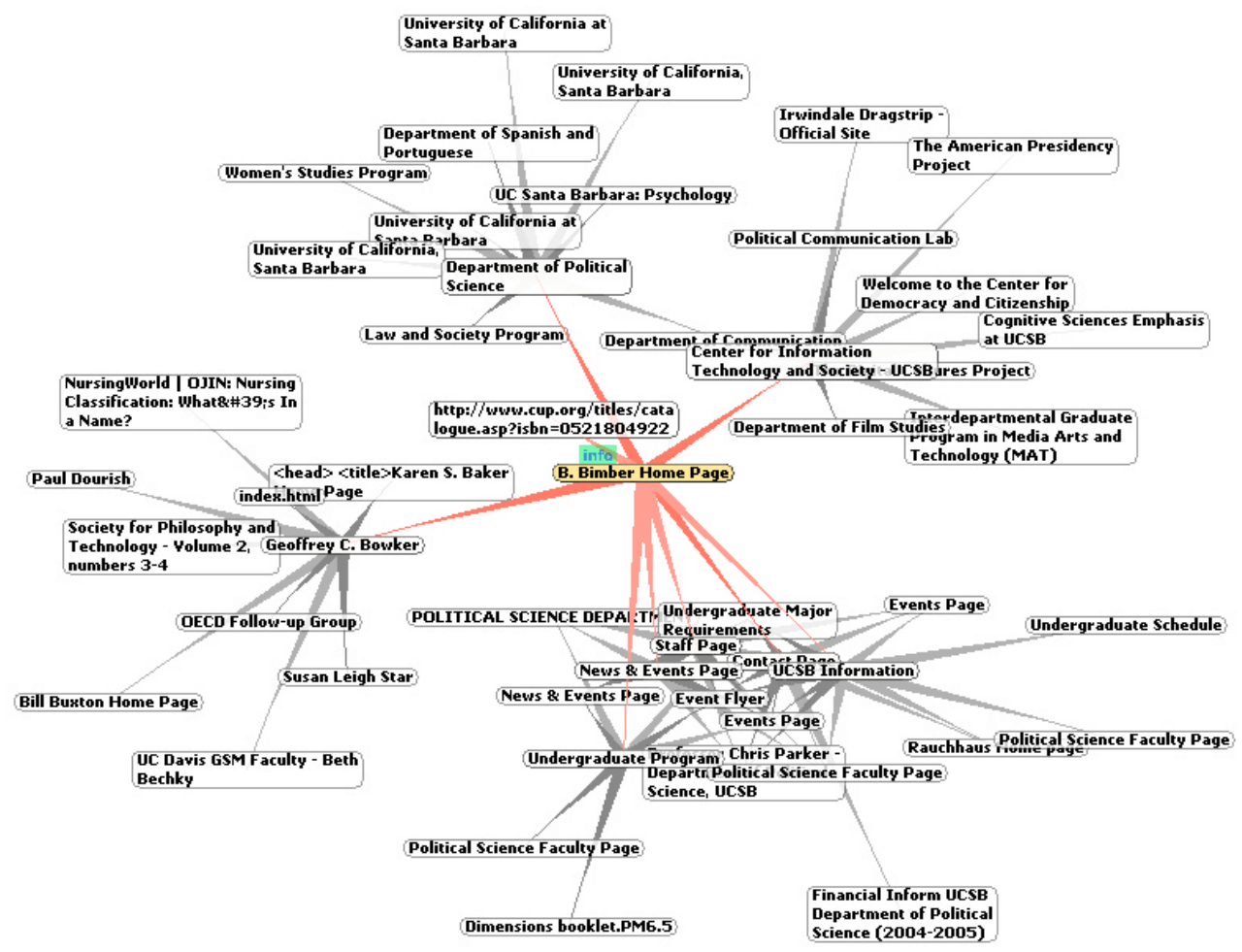
2. Information Topologies

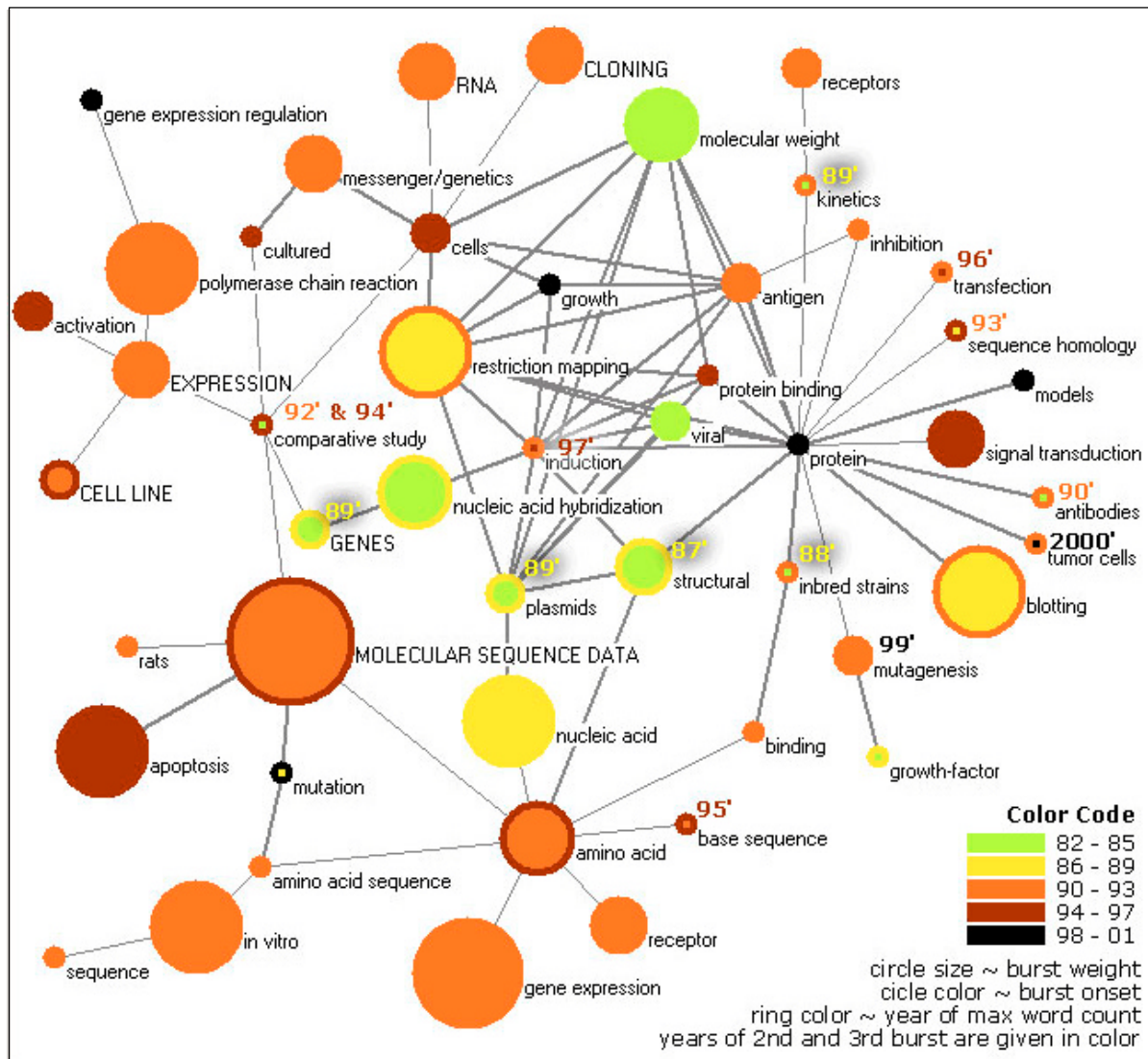
- the missing 'up button' on the browser
- intelligent summarisation and generalisation
- 3 key advantages:
 - a sense of the whole (the 'birds eye view')
 - revealing hidden connections
 - support interactive, unstructured browsing

Where is the wisdom we have lost in knowledge?

Where is the knowledge that we have lost in information.

T.S. Elliot, *The Rock* (1934)





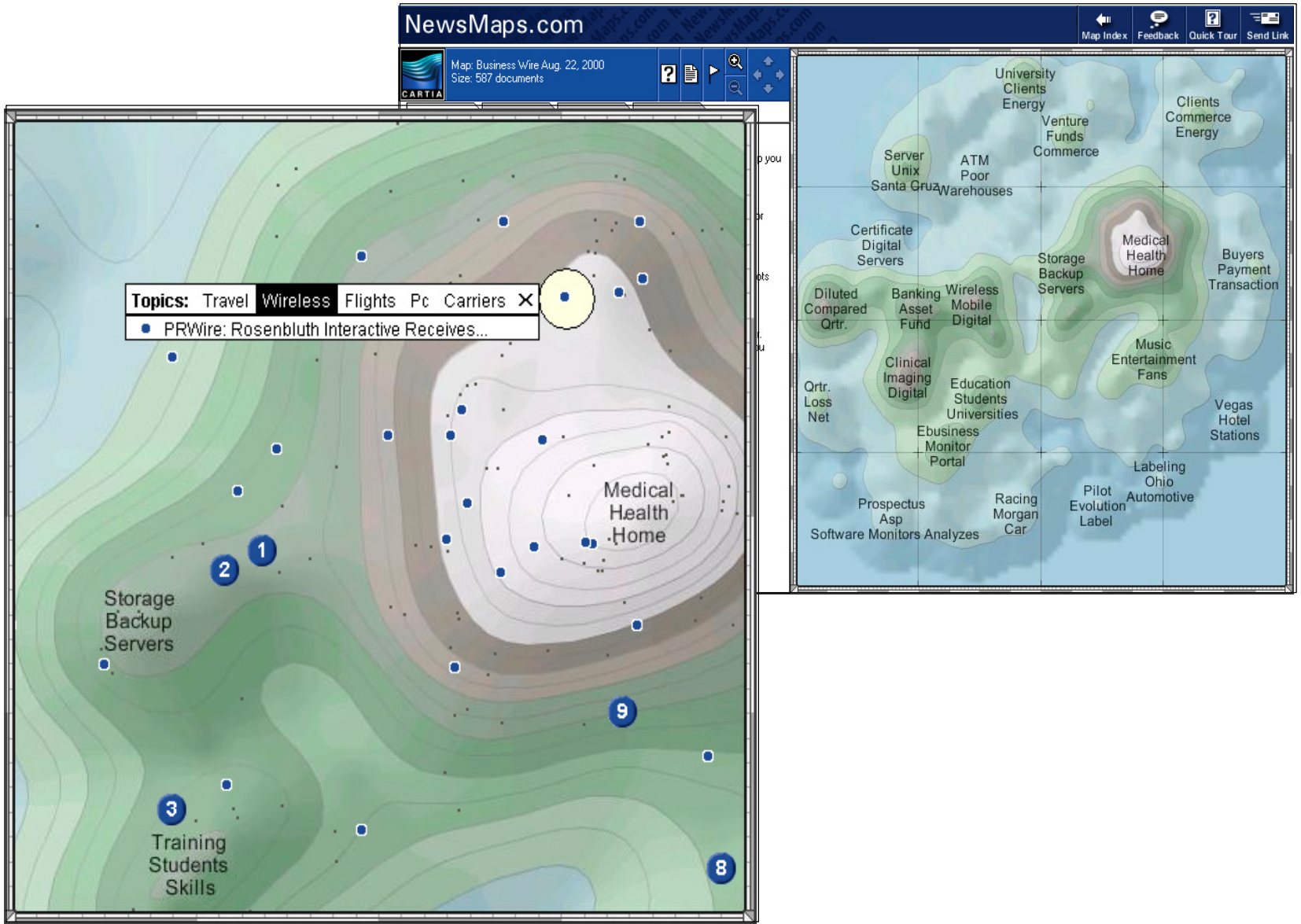
"A map of the top 50 "hot" words in the most highly cited PNAS articles from 1982-2001. Words appearing more often have larger circles, while the circle color and ring color identify when the word first appeared and when its popularity peaked, respectively."

- who makes 'the' map of science? all maps are authored frames
- automation of judgement. academic performance being sorted and ranked by 'universal' / 'objective' metrics based on highly detailed and individualised data
- similar across many social domains
- mapping facilitates 'social sorting', more effectively identifying 'good' and 'bad' risks
- many (most) in this room are likely to benefit. winners win more and others are more easily and effectively excluded

Spatialisation to make info-maps


- turning lot of non-geographic data into maps
- the application of explicit spatial metaphors as 'sense-making' tools for abstract data
- distance and location are based on some metric of semantic meanings - similarity of content/concepts
- represented using cartographic properties:
 - area
 - position
 - proximity
 - scale
 - + graphic properties of colour, shape, labels, etc

Cartia's ThemeScape - NewsMaps demo




NewsMaps.com


[Map Index](#)
[Feedback](#)
[Quick Tour](#)
[Send Link](#)



Map: Kosovo News 6 15 99
 Size: 951 documents

[Map Legend](#)
[Search](#)
[Topic List](#)
[Flags](#)

IP Vision Optical
Labels and Peaks
 Peaks represent large numbers of similar documents. Labels help you see what kind of information is contained in these areas.

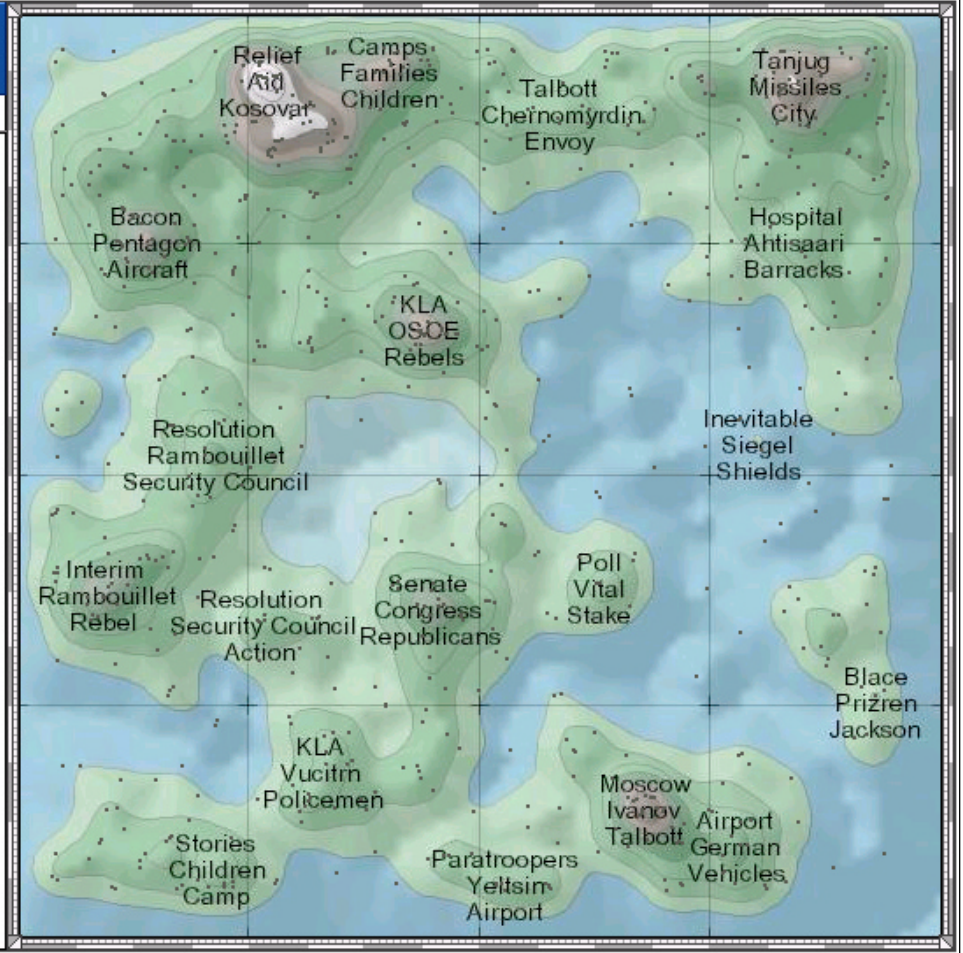

Document Points
 Dots represent documents on the map. Pause your mouse cursor over the map to view the topics in an area, or click to browse document titles.


Search Results
 Search results are drawn as blue dots on the map. Numbered dots represent the top documents displayed in the results list.


Flags
 Mark interesting documents with flags to help you find them later. You can also create summary pages listing all the documents you marked.

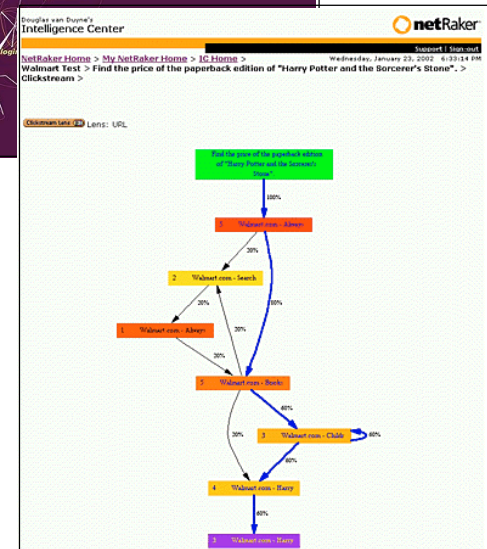
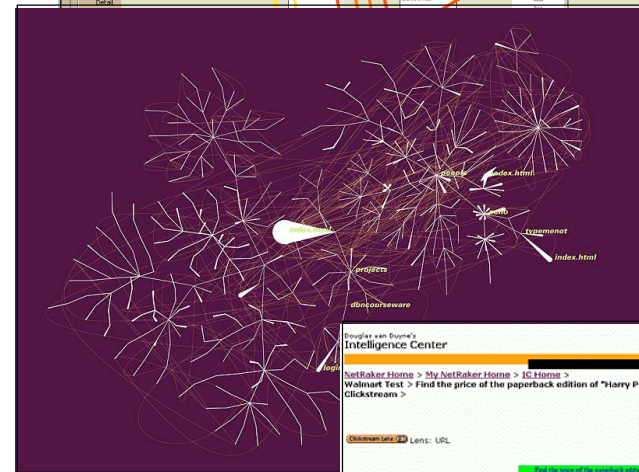
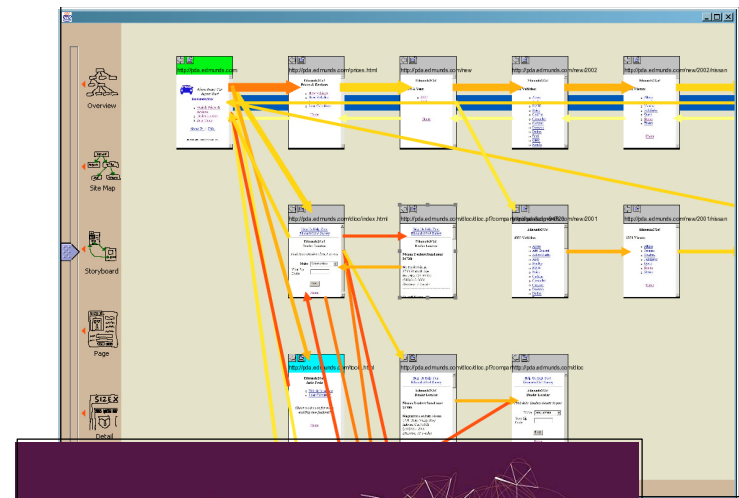
For more information:

[Quick Tour](#)
[Online Help](#)



3. Structures of online activities and interactions

- research
 - NIST's VISVIP
 - Stanford's WebQuilt
 - Ben Fry's Anemone
- commercial
 - Blue Martini's ClickViz
 - Insight's eBizinsights
 - Vividence ClickStreams
 - NetRaker Clickstream



Blue Martini's ClickViz

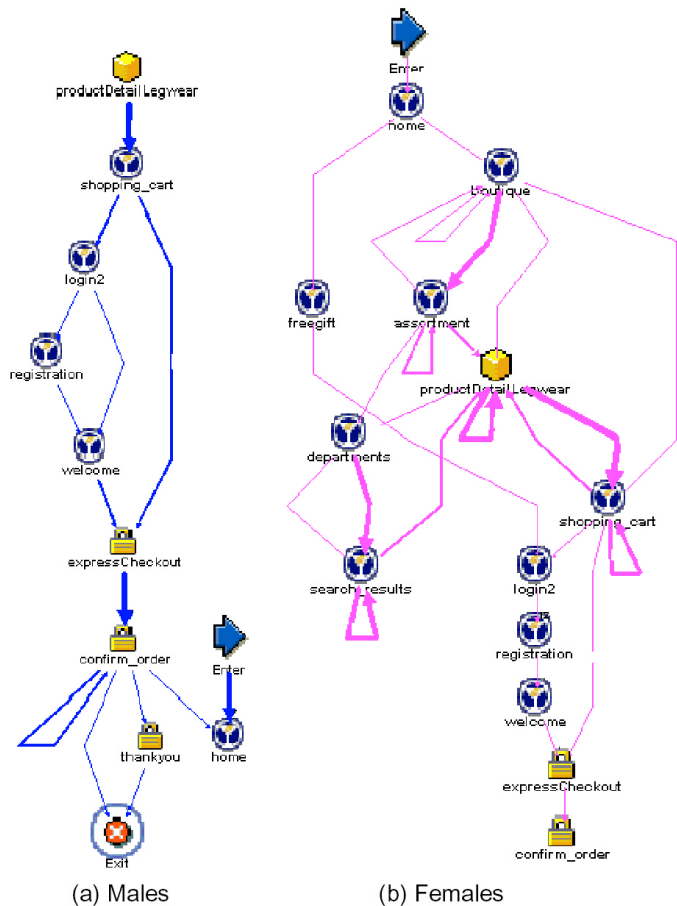


Figure 3. Gender Differences: Males tend to navigate in specific, direct patterns, whereas women's navigation patterns include much more browsing, utilizing much more of the site.

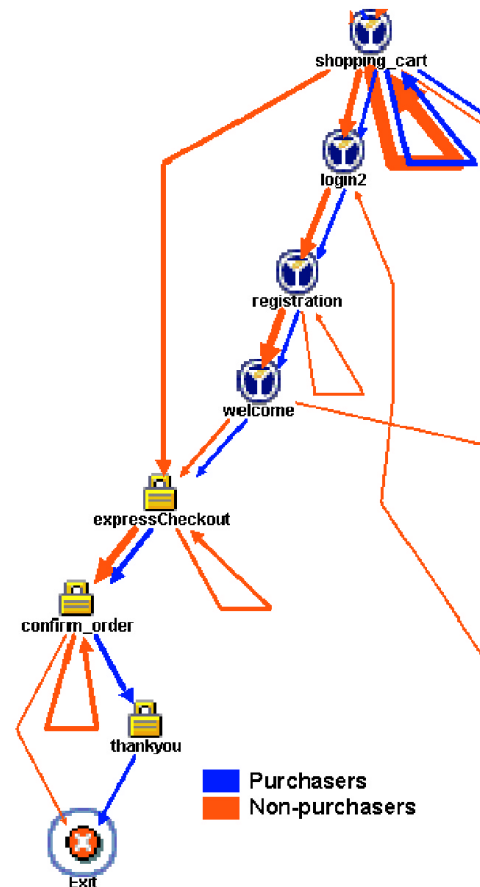
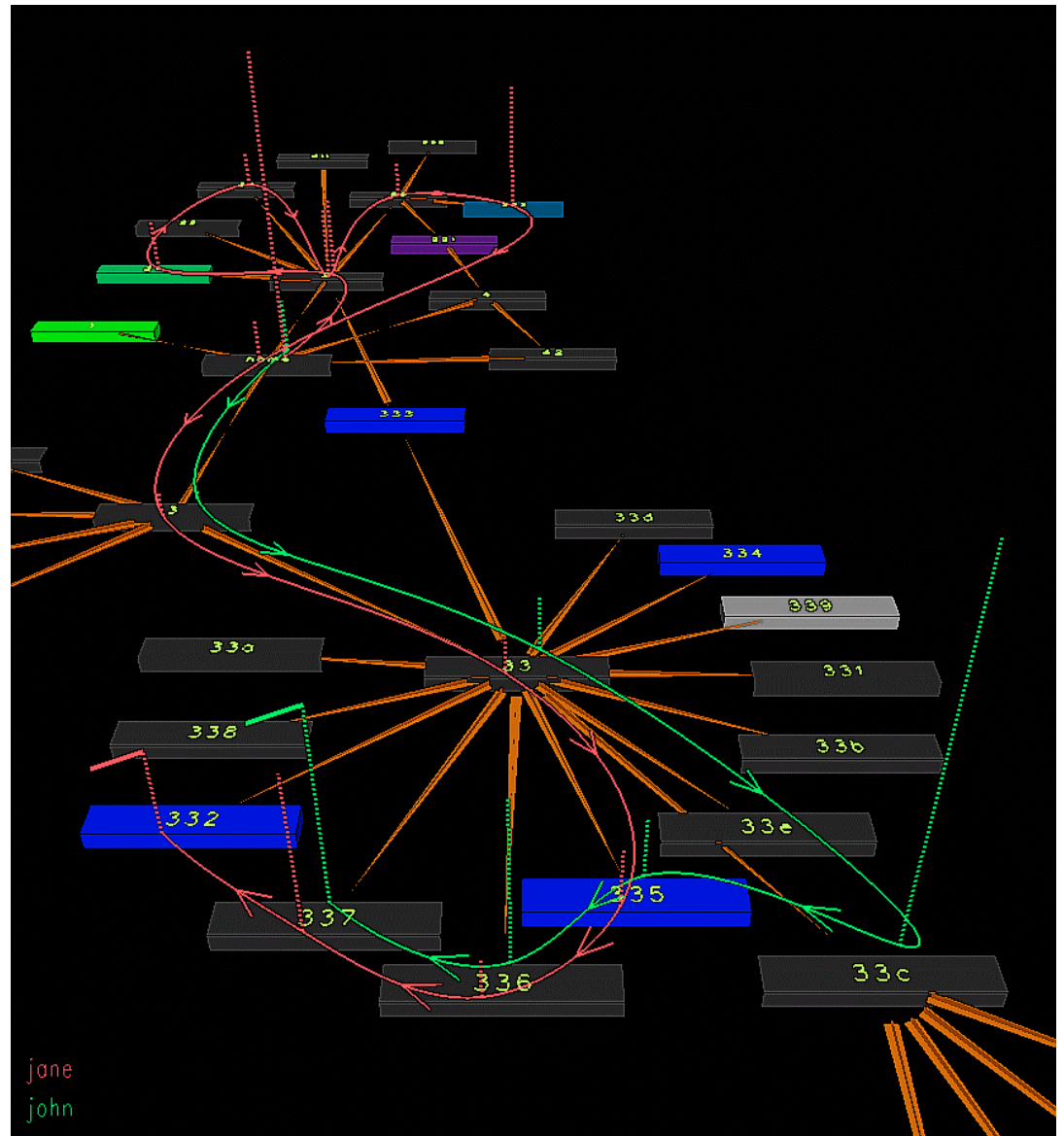


Figure 4. Checkout process. Purchasers take a direct route through the checkout process, whereas non-purchasers show a more haphazard route, including self-edges and early abandonment, possibly indicating a confusing checkout process.

VISVIP

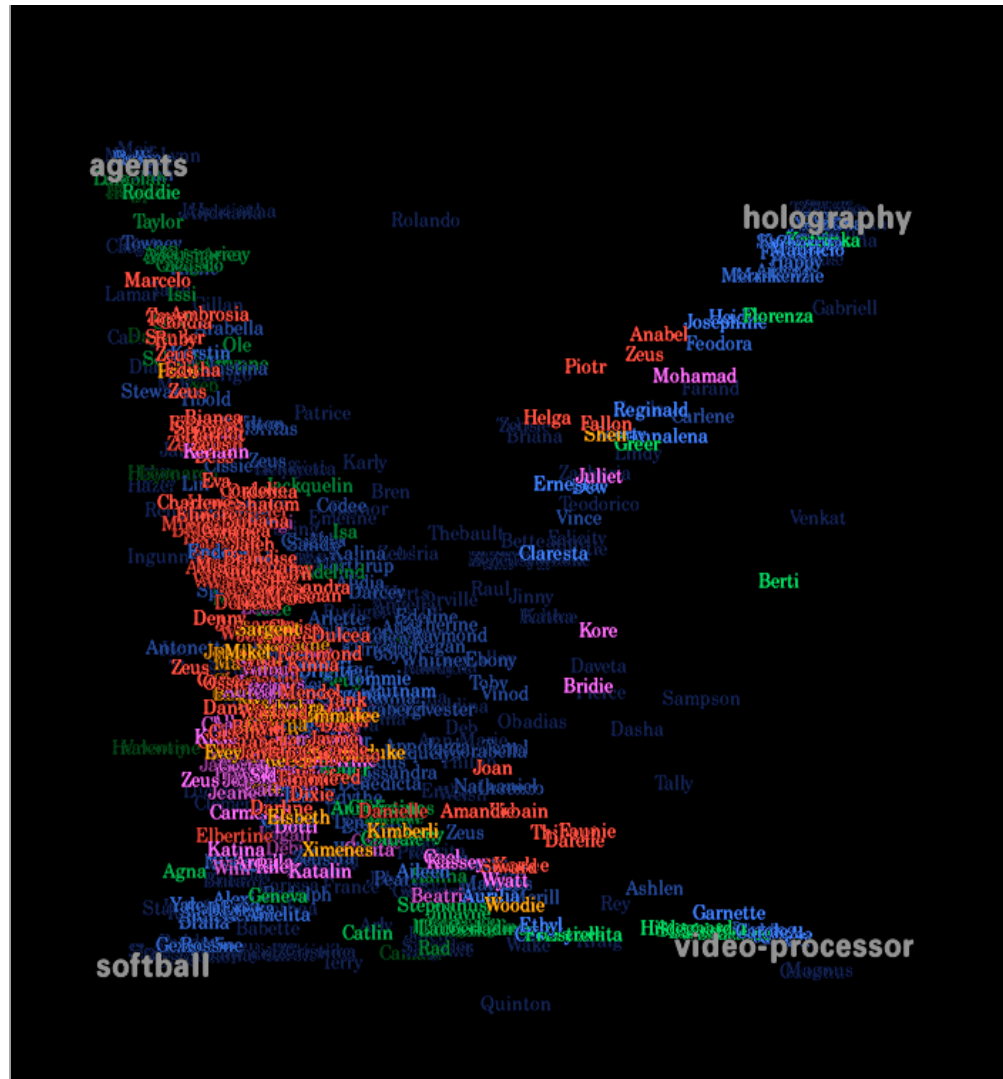
by John Cugini



(Source: <http://zing.ncsl.nist.gov/WebTools/VisVIP/overview.html>)

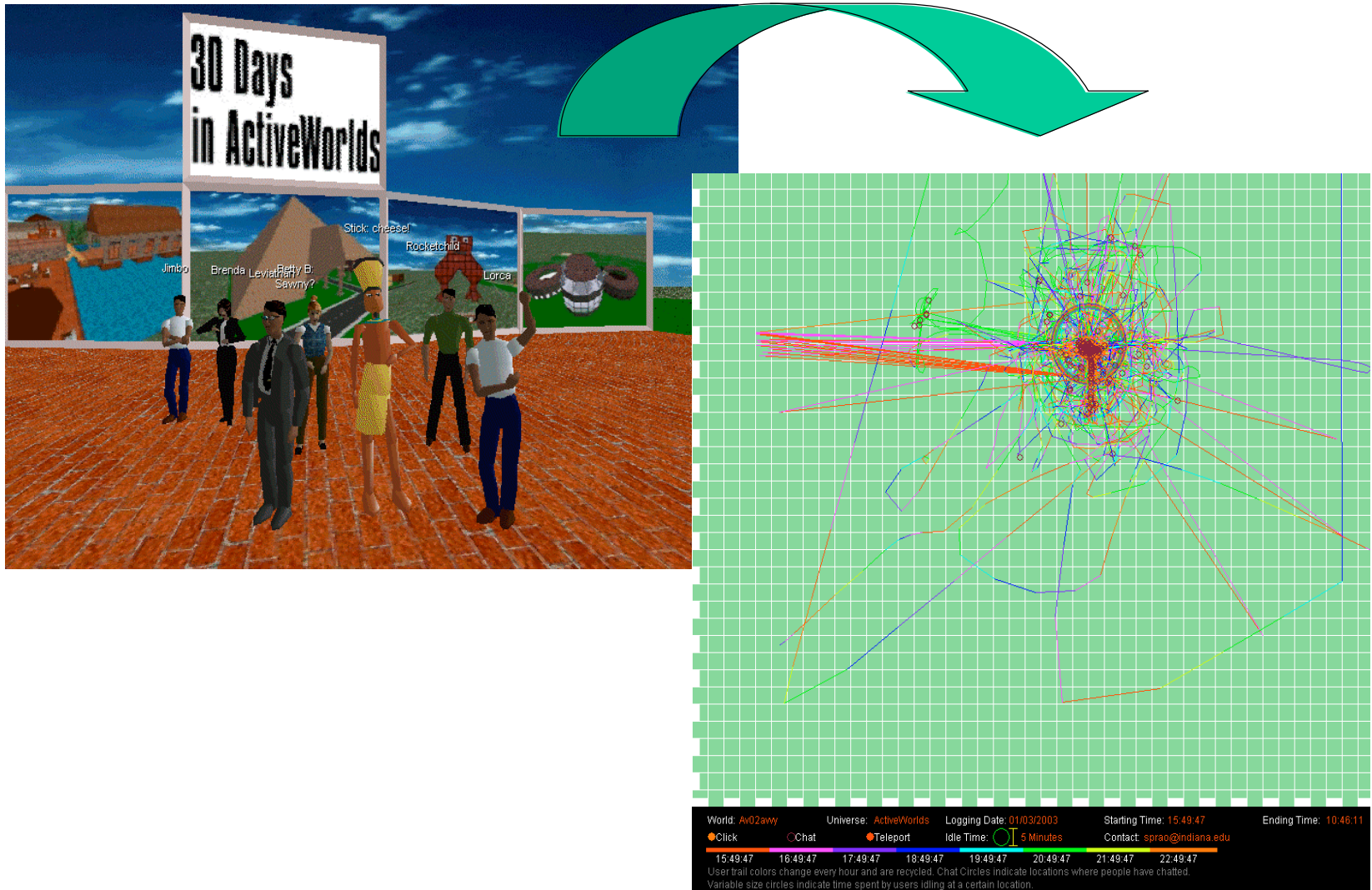
Mailing list affinities

VisualWho by
Judith Donath, MIT
Media Lab

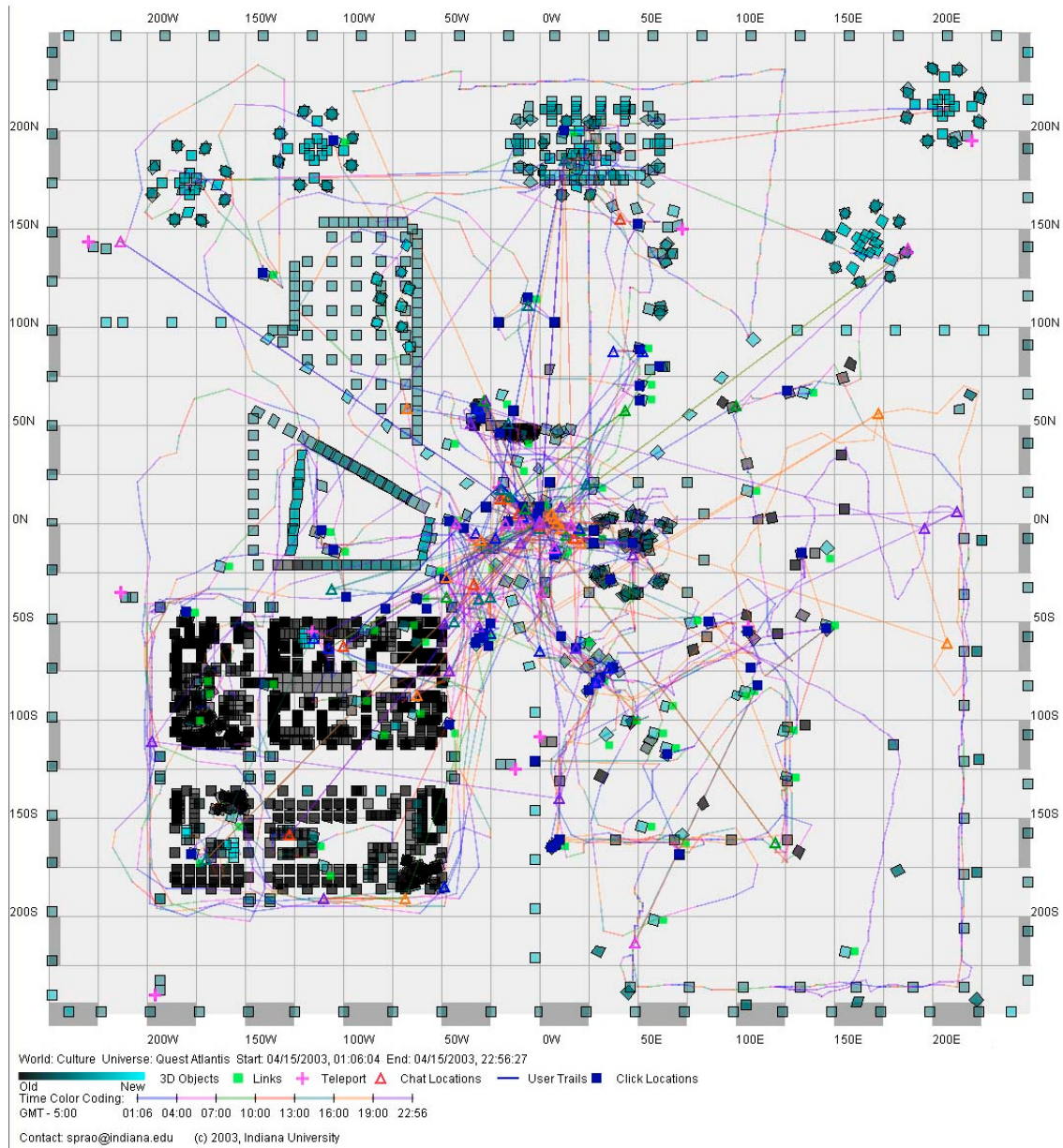


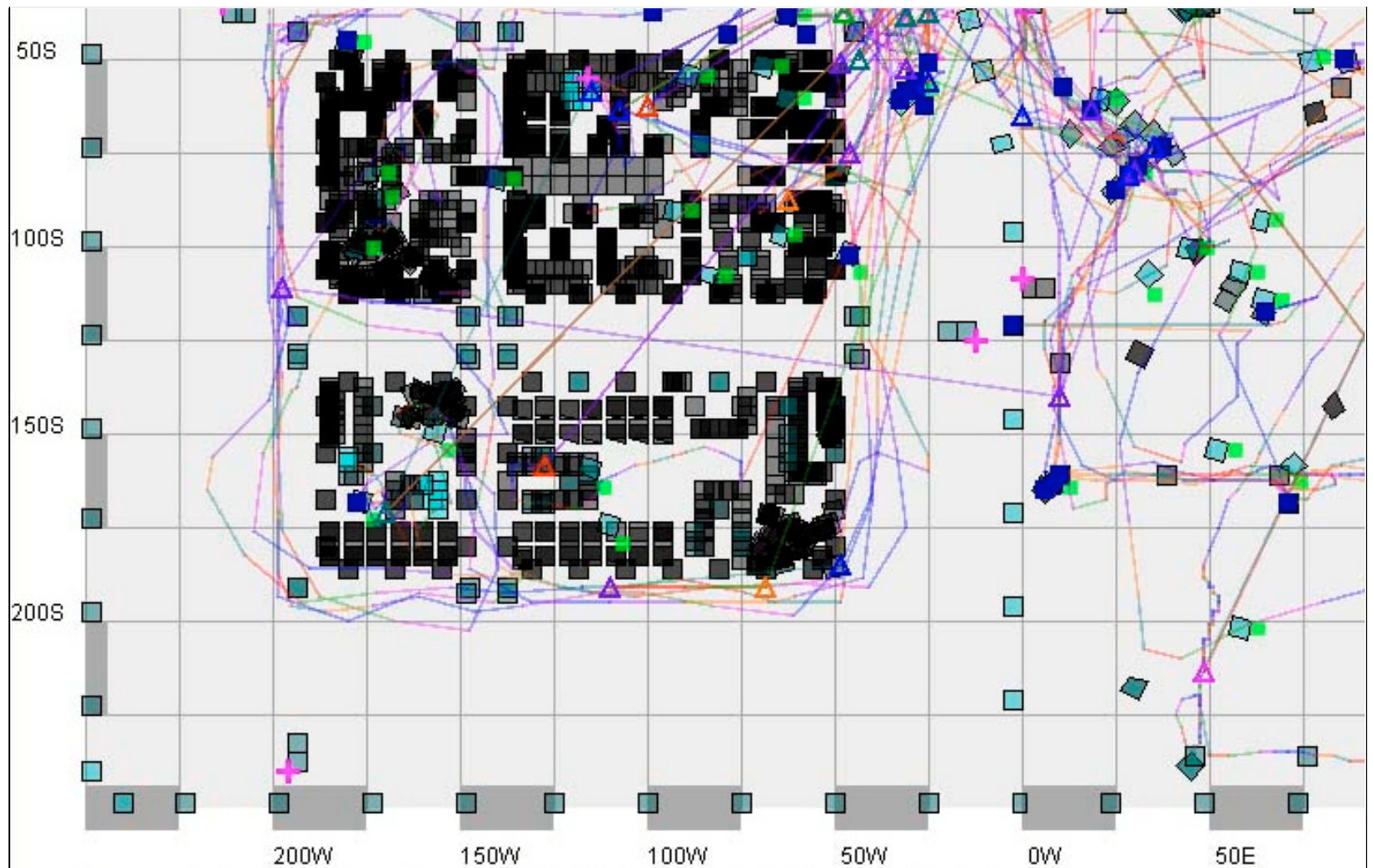
(Source: <http://smg.www.media.mit.edu/projects/VisualWho/>)

Individual 'movement' and activities in 3d virtual worlds



Interactive visualisation tool for analysing user trails in ActiveWorld space.





World: Culture Universe: Quest Atlantis Start: 04/15/2003, 01:06:04 End: 04/15/2003, 22:56:27

3D Objects
 Links
 Teleport
 Chat Locations
 User Trails
 Click Locations

Time Color Coding:
 GMT - 5:00 01:06 04:00 07:00 10:00 13:00 16:00 19:00 22:56

Contact: sprao@indiana.edu (c) 2003, Indiana University

Watching the Web

BBC NEWS UK EDITION

Last Updated: Thursday, 29 April, 2004, 10:36 GMT 11:36 UK

E-mail this to a friend Printable version

Spying software watches you work

Spyware has infected almost all companies polled for a survey about web-using habits at work.



Nine out of 10 of the technology managers questioned said machines at their firm had programs that spied on the browsing habits of staff.

The computer staff estimated that, on average, 29% of work PCs had spyware surreptitiously installed on them.

By contrast only 6% of users questioned believed that the machine they use had been infected by such software.

Browsing risk

Have Your Say Magazine In Pictures Week at a Glance Country Profiles In Depth Programmes

BBC NEWS UK EDITION

Last Updated: Monday, 5 April, 2004, 22:27 GMT 23:27 UK

E-mail this to a friend Printable version

Google's Gmail sparks privacy row

Internet search engine Google's plans for a free email service have come under fire from privacy campaigners.



Google is devising Gmail as a rival to Microsoft's Hotmail and to Yahoo!


Privacy campaigners have objected to plans to send users adverts linked to the content of messages, and to the permanent storage of email.



OBSTACLES TO THE FREE FLOW OF INFORMATION ONLINE

THE INTERNET UNDER SURVEILLANCE

REPORTERS WITHOUT BORDERS 2003 REPORT



CNN.com MEMBER SERVICES International Edition | Netscape

MAKE CNN.com YOUR HOME PAGE

SEARCH The Web CNN.com Search Powered by **YAHOO!** search

TECHNOLOGY

Google bows to Chinese censorship

Monday, September 27, 2004 Posted: 9:54 AM EDT (1354 GMT)

SAN FRANCISCO, California (AP) -- Google Inc.'s recently launched news service in China doesn't display results from Web sites blocked by that country's authorities, raising prickly questions for an online search engine that has famously promised to "do no evil."

Dynamic Internet Technology Inc., a research firm striving to defeat online censorship, conducted tests that found Google omits results from the government-banned sites if search requests are made through computers connecting to the Internet in China.

Steered by an identical search request, computers with a United States connection retrieved results from the sites blocked by China.

"That's a problem because the Chinese people need to know there are alternative opinions from the Chinese government and there are many things being covered up by the government," said Bill Xia, Dynamic's chief executive. "Users expect

advertiser links what's this?

Refinance Rates Just Got Lower
Lock in the lowest rates of the summer. Get up to 4 free refinance quotes.
www.lowerrmybilis.com

LendingTree.com - Official Site
Lendingtree - Find a mortgage, refinance, home equity or auto loan now. Receive...
www.lendingtree.com

Hundreds of Online Degree Programs
Search our directory of online degree programs. Associate, Bachelor, Master and...

Search Jobs more options

Enter Keywords
Enter City ALL SEARCH

careerbuilder.com

Level (3) COMMUNICATIONS
The Network Partner You Can Rely On.™

Commerce likes to know you (a lot)

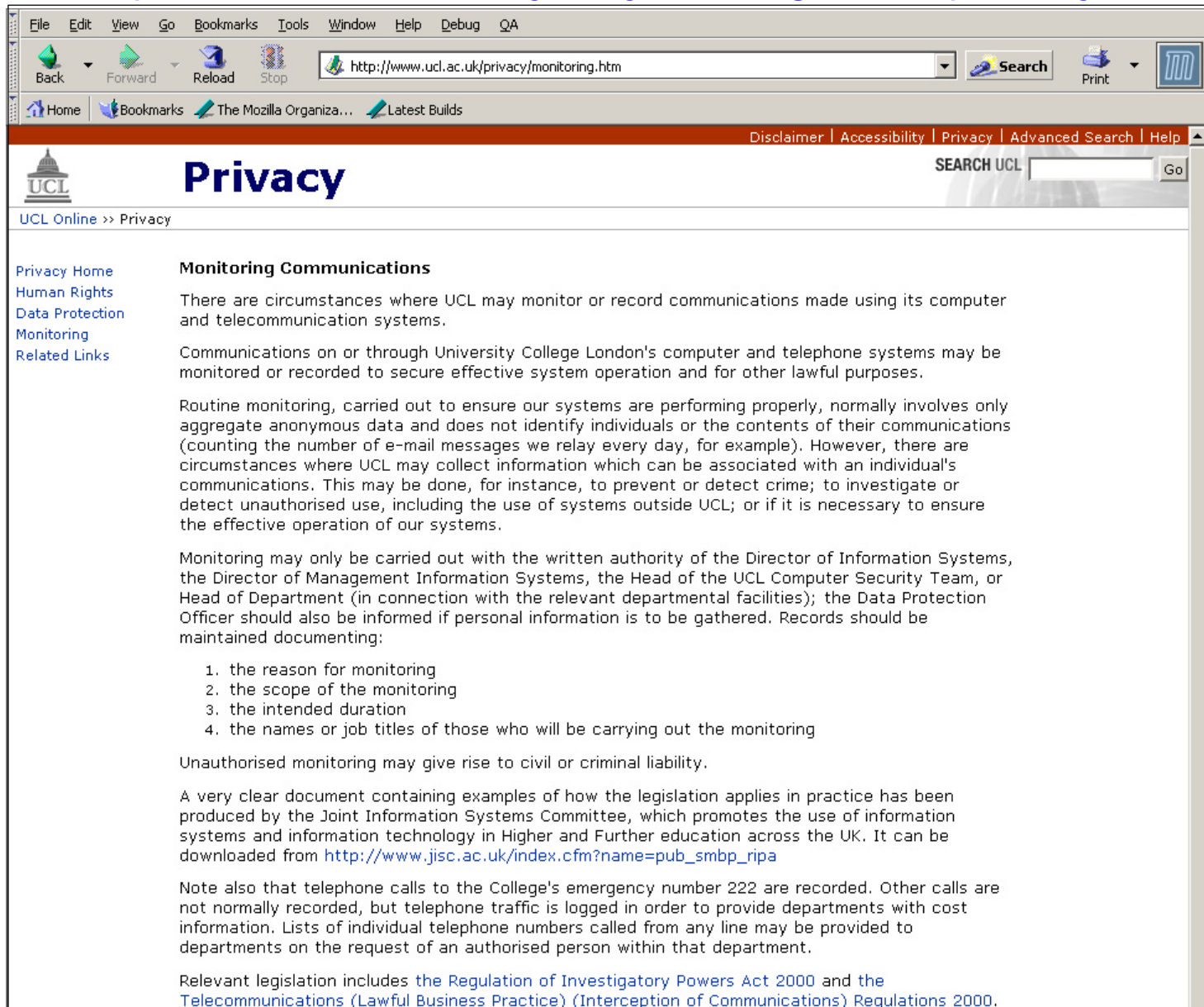
According to Ron Kohavi director of data mining and personalization at Amazon.com

E-mail Campaigns



- **When doing an e-mail campaign, make sure to**
 - **Personalize the e-mail**
 - **Make every link unique (redirection link) to allow for user identification on clickthrough**
 - **Track revenues by clickthroughs and visits by recipients to site following a campaign**
 - **Use unique coupon codes to identify user across all channels (e.g., use of e-coupon at stores)**
 - **Put a “web bug” (single-dot image) that is retrieved from the server on e-mail opening**
Allows computing the e-mail “open rate”

Expectations of anonymity and rights to privacy?



The screenshot shows a Mozilla browser window with the address bar containing <http://www.ucl.ac.uk/privacy/monitoring.htm>. The page title is "Privacy" and the UCL logo is visible in the top left. A navigation menu on the left includes links for "Privacy Home", "Human Rights", "Data Protection", "Monitoring", and "Related Links". The main content area is titled "Monitoring Communications" and contains the following text:

There are circumstances where UCL may monitor or record communications made using its computer and telecommunication systems.

Communications on or through University College London's computer and telephone systems may be monitored or recorded to secure effective system operation and for other lawful purposes.

Routine monitoring, carried out to ensure our systems are performing properly, normally involves only aggregate anonymous data and does not identify individuals or the contents of their communications (counting the number of e-mail messages we relay every day, for example). However, there are circumstances where UCL may collect information which can be associated with an individual's communications. This may be done, for instance, to prevent or detect crime; to investigate or detect unauthorised use, including the use of systems outside UCL; or if it is necessary to ensure the effective operation of our systems.

Monitoring may only be carried out with the written authority of the Director of Information Systems, the Director of Management Information Systems, the Head of the UCL Computer Security Team, or Head of Department (in connection with the relevant departmental facilities); the Data Protection Officer should also be informed if personal information is to be gathered. Records should be maintained documenting:

1. the reason for monitoring
2. the scope of the monitoring
3. the intended duration
4. the names or job titles of those who will be carrying out the monitoring

Unauthorised monitoring may give rise to civil or criminal liability.

A very clear document containing examples of how the legislation applies in practice has been produced by the Joint Information Systems Committee, which promotes the use of information systems and information technology in Higher and Further education across the UK. It can be downloaded from http://www.jisc.ac.uk/index.cfm?name=pub_smbp_ripa

Note also that telephone calls to the College's emergency number 222 are recorded. Other calls are not normally recorded, but telephone traffic is logged in order to provide departments with cost information. Lists of individual telephone numbers called from any line may be provided to departments on the request of an authorised person within that department.

Relevant legislation includes [the Regulation of Investigatory Powers Act 2000](#) and [the Telecommunications \(Lawful Business Practice\) \(Interception of Communications\) Regulations 2000](#).

Privacy through obscurity

- liberties come, in large part, because of the inherent administrative inefficiencies in surveillance
- rapid developments in data mining of online interactions are changing the balance:
 - looking deeper, easy to zero in on a single person
 - looking backwards. many more things kept and getting kept for longer time
 - profiles across time and space. prediction of future
- increasingly permeable borders to private information. unintended leakage of personal data
- widening access. 'bottom-up' surveillance. anyone can google me. (e.g. Google access to Usenet archives; Way Back Machine at Internet Archive)

Maps make a difference?

- rise of visual data mining for e-research combining powerful capabilities of software and people, exploits best abilities of both (fast but dumb logging/analysis plus slow but smart pattern recognition and interpretation)
- maps “...hold out the promise that the user will be able not only to visualize a few nearby trees in the forest of knowledge, but also to understand the entire landscape. If these techniques can be made to operate effectively, they may well change the way that science is conducted and the way the business of the world is carried out.” (Richard M. Shiffrin and Katy Borner, PNAS, 2004)
- effective maps in the e-research context reveal novel insights that are not apparent with other forms of analysis

Ethics of e-researchers and their tools

- maps don't kill people
- maps of social cyberspaces are 'responsible artefacts', that do not destroy what they seek to represent or enhance
- "The bright light of social science research can create an unpleasant glare for participants drawn to a dimly lit online space." (Marc Smith, 1998, "Invisible crowds in cyberspace: Mapping the social structure of the Usenet", in *Communities in Cyberspace*)
- maps of e-research can prove to be very valuable, but at the same time they can never be value-free

- questions?
- these slides will be at www.casa.ucl.ac.uk/martin/canberra.pdf
- more maps at www.cybergeography.org/atlas
- email, m.dodge@ucl.ac.uk

Bibliography:

- Terra Australis Incognita map courtesy of the *National Library of Australia*, <http://nla.gov.au/nla.map-nk2456-13>
- U.S. telegraph map courtesy of the *Library of Congress*, <http://hdl.loc.gov/loc.gmd/g3701p.ct000084>
- movie of Code Red worm spread courtesy of *CAIDA*, www.caida.org/analysis/security/code-red/
- Katy Börner work, Indiana University <http://ella.slis.indiana.edu/~katy/>
- special issue of *PNAS* on 'mapping knowledge domains', April 6th, 2003, vol.101
- Lumeta *Internet Mapping Project*, www.lumeta.com/mapping.html