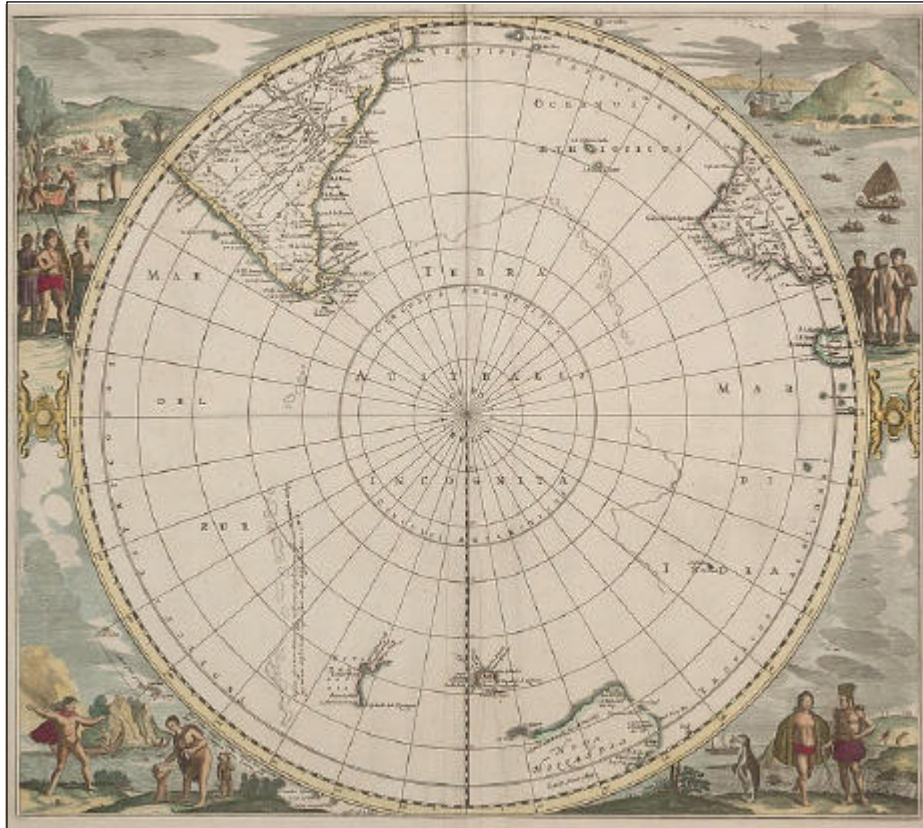# Cartography, privacy and surveillance: Are some parts of the Web better left unmapped?

Martin Dodge

Centre for Advanced Spatial Analysis, University College London

Web: espaces de navigation, objets d'exploration, 13th May 2004

"… we have every reason to believe that cyberspace, left to itself, will not fulfill the promise of freedom. Left to itself, cyberspace will become a perfect tool of control … The invisible hand of cyberspace is building an architecture that is quite the opposite of what it was at cyberspace's birth."

Lawrence Lessig, *Code and Other Laws of Cyberspace*

Last Updated: Thursday, 26 February, 2004, 16:49 GMT

✉ E-mail this to a friend   🖨 Printable version

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Politics
Health
Education
Science/Nature
Technology
Entertainment

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

BBC SPORT
BBC WEATHER
CBBC news
BBC ON THIS DAY

UK 'spied on UN's Kofi Annan'

British spies listened in to UN Secretary General Kofi Annan's office in the run up to the Iraq war, former UK cabinet minister Clare Short says.

Ms Short says she expects there are transcripts of calls to Mr Annan

Ms Short said she had read transcripts of some of Mr Annan's conversations.

She said she recalled thinking, as she talked to Mr Annan: "Oh dear, there will be a transcript of this and people will see what he and I are saying."

Tony Blair said the claims were "deeply irresponsible" and appeared to cast doubt on her future as a Labour MP.

Short's future?

UN officials said they did not know whether the allegations were true or not, but say such actions would have been illegal.

The secretary general's spokesman, Fred Eckhardt, said [...] Annan had nothing to hide and that anyone who wanted [...] know his opinion on an issue just had to ask him to his fa[...] [...] e the prime minister was [...] ort's comments.

The University and its constituent Colleges operate 24 hour CCTV surveillance

Images are being recorded for the purposes of crime prevention and detection. Please contact 020 7862 8133 for further information about the scheme.

This scheme is controlled by: The University of London, Senate House, Malet Street, London, WC1E 7HU.

UCL   Birkbeck UNIVERSITY OF LONDON

CCTV IN OPERATION

WESTMINSTER CCTV

Making our streets safer

City of Westminster   METROPOLITAN POLICE

TEL. 020 7641 6511

   ⏯ WATCH/LISTEN TO BBC NEWS

Last Updated: Saturday, 24 April, 2004, 07:11 GMT 08:11 UK

✉ E-mail this to a friend   🖨 Printable version

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Politics
Health
Education
Science/Nature
Technology
Entertainment

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

BBC SPORT
BBC WEATHER
CBBC news
BBC ON THIS DAY

ID card trials to start next week

Trials of identity cards are to be launched next week, the BBC has learnt.

The pilot will involve 10,000 volunteers and be run from the Passport Office in London and three other centres around Britain.

The government hopes the pilot scheme will pave the way for compulsory identity cards for everyone within the next decade.

An opinion poll suggests most people back an ID scheme

Ministers are due to set out details of plans for a nationwide identity database on Monday.

Biometric options

They will publish draft legislation and Home Secretary David Blunkett says he wants a Bill paving the way for the scheme to be passed before the next general election.

Carrying false identity papers is also to be made a specific offence for the first time under the plans, with offenders facing up to 10 years in jail, say government sources.

WATCH AND LISTEN

The BBC's Shaun Ley
"Proving the proposed cards can't be faked is an important stage in winning over doubters"

▶ VIDEO

BBC iCAN

Identity cards
Find out more on the key issues and what you can do

SPEC[...]

LABOUR'S ID CARD PLANS

LATEST STORIES
▸ ID card trials to start
▸ Public 'happy to carry ID cards'
▸ ID cards have 'hidden dangers'
▸ Met's chief backs identity cards
▸ ID cards 'could worsen racism'

Will they work?
The pros and cons of ID cards

ANALYSIS & BACKGROUND

WATCHING THE WEB

BBCNEWS 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.

Spyware watches you from a distance

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

Perfect panopticon for commerce?

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

The Web: Good

BLUE MARTINI

- **The web is a special channel with huge advantages for mining data**
  - **"Perfect" data collection is possible**
    - Every interaction (page view, search, form) can be recorded
    - Views and transactions (e.g., purchases) are easily correlated
    - A lot of data is available quickly
      - Even a small site selling 5 items an hour will have 1.6 million page views after the first month
  - **Electronic collection more reliable other channels where data is entered by hand**
  - **Actionable – easy to change things online**
- **The web is also a place for customers to do research.** Forrester claims that "29% of the online population researched purchases online to buy them offline" and that the web "will influence 26% of total retail sales" by 2005

5

(source: http://robotics.stanford.edu/users/ronnyk/etailKohavi.pdf)

Defining privacy

- Justice Brandeis (1890), '*the right to be let alone*'

- modern privacy benchmark, at an international level, can be found in the 1948 Universal Declaration of Human Rights, which specifically protects territorial and communications privacy. Article 12 states:

- '*No one should be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks on his honour or reputation. Everyone has the right to the protection of the law against such interference or attacks.*' (www.un.org/Overview/rights.html)

- the power to control what others can come to know about you

Web and individual privacy

- nothing new. ongoing tension and debate
- most express concern but actions are ambivalent

BBCNEWS UK EDITION

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Market Data
Your Money
E-Commerce
Economy
Companies
Fact Files
Politics
Health
Education
Science/Nature
Technology
Entertainment

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

BBC SPORT

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.

Campaign group Privacy International has filed a complaint with the UK Information Commissioner, Reuters news agency reported.

The Information Commissioner's job is to make sure organisations comply with a web of laws safeguarding privacy and freedom of information.

Online snooping?
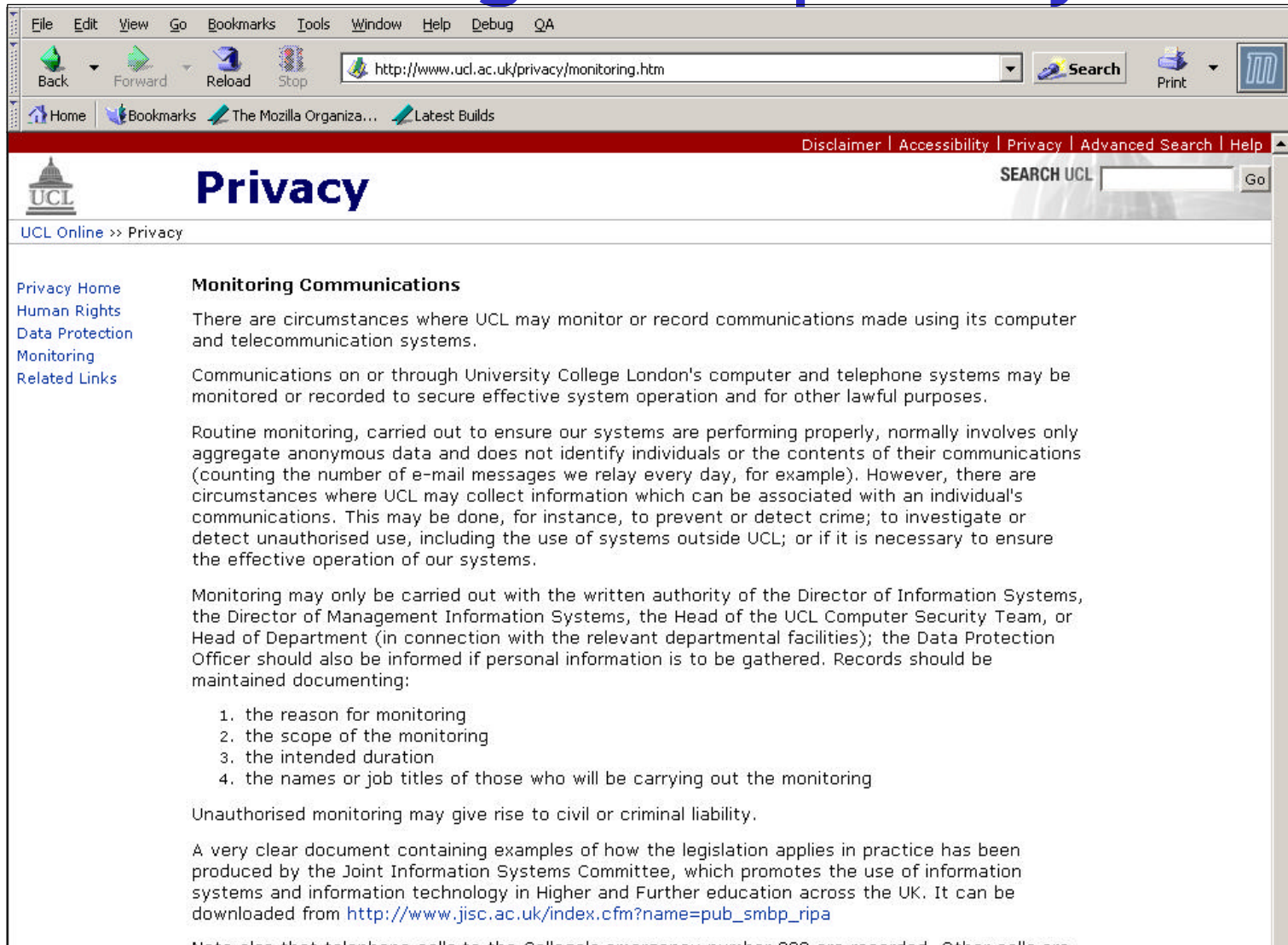
Privacy International objects to Google's plans to send users links to advertising based on a computer scan of their correspondence, and presumed interests.

Expectations of anonymity and rights to privacy?

File Edit View Go Bookmarks Tools Window Help Debug QA

Back Forward Reload Stop http://www.ucl.ac.uk/privacy/monitoring.htm Search Print

Home Bookmarks The Mozilla Organiza... Latest Builds

Disclaimer | Accessibility | Privacy | Advanced Search | Help

Privacy

SEARCH UCL Go

UCL Online >> Privacy

Privacy Home
Human Rights
Data Protection
Monitoring
Related Links

Monitoring Communications

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

Leaky browsing

Basic Information

Shows some of the basic information available in the browser.

Variable	Setting
appName	Microsoft Internet Explorer
appCodeName	Mozilla
appVersion	4.0 (compatible; MSIE 6.0; Windows NT 5.0)
appMinorVersion	; SP1; Q810847; Q813951;
IE real version	6.0 - 6,0,2800,1106
UserAgent	Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)
Security Policy	Property not supported or blank!
Browser Name	Internet Explorer 6
Operating System	Windows - Windows 2000
Platform	Win32 - Windows 32-bit platform
CPU	x86 - x86 Compatible processor (Intel, Cyrix, AMD, etc.)
Online	true - The system is in g
Connection Type	modem - You are connec

Date Information

Time and date information from your system.

Variable	Setting
Time and date	Mon May 3 19:28:47 UTC+0100 2004
GMT'ed	Mon, 3 May 2004 18:28:47 UTC
Locale String	03 May 2004 19:28:47
Time-Zone offset	-60 minutes
Milliseconds since January 1, 1970, 00:00:00	1083608927000
Internet Beat	@811

'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 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
- unintended leakage, permeable borders to private information
- widening access. 'bottom-up' surveillance. anyone can google me. (e.g. Google access to Usenet archives; Way Back Machine at Internet Archive)

Googledorks!

johnny.ihackstuff.com

"I'm johnny. I hack stuff."

Welcome, guest! Please **create an account**   (what's here?) downloads forum! googledorks! photos/art web links search   _MAY 10, 2004

Latest Googledorks

#212: "Request Details" "Control Tree" "Server Variables"
#211: "HTTP_FROM=googlebot" googlebot.com "Server_Software="
#209: filetype:conf inurl:firewall -intitle:cvs
#210: filetype:inc intext:mysql_connect
#208: "index of" inurl:recycler
#207: inurl:names.nsf? opendatabase
#206: filetype:properties inurl:db intext:password
#205: intitle:"Samba Web Administration Tool" intext:"Help Workgroup"
#204: inurl:"smb.conf" intext:"workgroup" filetype:conf conf
#203: filetype:pdf "Assessment Report" nessus

Recent Files

...more files...

goolink.pl added 05/04/04

The Google Hacker's Guide added 03/19/04

Gooscan Google Scanner added 03/11/04

15 Minute Security Guide: SSH added 02/01/04

web hacking walkthrough (doc) added 09/22/03

Googledorks!

googleDork (gOO gÃ´l'DÃ´rk) noun 1. Slang. An inept or foolish person as revealed by Google.

Entries by Category

Error Messages (36)
Really retarded error messages that say WAY too much!

Files containing juicy info (67)
No usernames or passwords, but interesting stuff none the less. *new*

Files containing passwords (30)
PASSWORDS, for the LOVE OF GOD!!! Google found PASSWORDS! *new*

Files containing usernames (8)
These files contain usernames, but no passwords... Still, google finding usernames on the web.. *new*

Footholds (1)
Examples of queries that can help gain a foothold into a web server

Pages containing login portals (10)
These are login pages for various services. Consider them the front door of a site's more sensitive functions. *new*

Pages containing network or vulnerability data (6)
These pages contain such things as firewall logs, network information, IDS logs... all sorts of fun stuff! *new*

Sensitive Directories (17)
Google's collection of sites sharing sensitive directories. The files contained in here will vary from sesitive to uber-secret! *new*

Sensitive Online Shopping Info (0)
Examples of queries that can reveal online shopping info like customer data, suppliers, orders, credit info, etc

Vulnerable Files (3)
HUNDREDS of vulnerable files that Google can find on web servers...

Vulnerable Servers (18)
These searches reveal servers with specific vulnerabilities. These are found in a different way than the searches found in the "Vulnerable Files" section. *new*

Web Server Detection (13)
These links demonstrate Google's awesome ability to profile web servers..

There are **212** entries and **12** Categories in our database

10 most popular entries   **10 most recent entries**

Benefits of forgetfulness ?

Google Groups

Web   Images   **Groups**   News   Froogle^New!   **more »**

"martin dodge"   [Search]   **Advanced Groups Search**
                                Preferences

Groups   Search result 49 for **"martin dodge"**

Accurate IP Intelligence ▪ Accurately identify Internet users to a city level, in real-time ▪ www.digitalenvoy.net   Sponsored Links
MaxMind IP to Location ▪ Free and Paid Datab...   ... API ▪ maxmind.com/
IP to Country Database ▪ For PHP, ASP, JSP a...   ...ad at $49 per year. ▪ ip2location.com

1996

From: **Martin Dodge** (mdodge@vulca...   Search Result 49
Subject: The Geography of Cyberspac...
Newsgroups: bit.listserv.geograph   View: Complete Thread (6 articles)
Date: 1996/11/06   Original Format

Hello,

Just to let any interested people know that my web page on "The
Geography of Cyberspace" has moved to a new WWW address. I have changed
jobs, so my web pages have had to be moved. The new URL of "The Geography
of Cyberspace" page is :

URL still works!

http://www.geog.ucl.ac.uk/casa/martin/geography_of_cyberspace.html

Please update any bookmark you may have. Sorry for the inconvenience.

If you don't know about my "Geography of Cyberspace" page, it is simply
a web page providing pointers to a range of information I have found
useful in exploring the geographies of the Internet, WWW and
cyberspace.

I would welcome any feedback or comments you might have on my web page.

Thanks for your time,
martin dodge

--

===
MARTIN DODGE - email : m.dodge@ucl.ac.uk
Centre for Advanced Spatial Analysis (CASA)
University College London,
1-19 Torrington Place,
London, WC1E 6BT. United Kingdom.

Tel : 0171 391 1782
Web : www.geog.ucl.ac.uk/casa/
===

Searched for **ht**

Note some duplicat
* denotes when sit

1996
0 pages

Jun
Oct

The Geography of Cyberspace

Last revised on 09/06/97. Maintained by **Martin Dodge** *(M.Dodge@ucl.ac.uk).*
Centre for Advanced Spatial Analysis (CASA), *University College London, United Kingdom.*

You are (approximately) visitor number `001043` *since 02/06/97.*

I am interested in understanding the nature and geography of Cyberspace. Here is a rather eclectic list of the useful resources I have come across. They relate to various resources for measuring and mapping the geography of an embryonic Cyberspace (i.e. the Internet and WWW).

If you would like to be kept informed of updates to this page please register to receive the 🌐 Geography of Cyberspace Update Bulletin 🌐. Any comments you may have are most welcome. If you would like to add a link please feel free to contact me.

What's New on The Geography of Cyberspace Page?

09/06/97:

- ISP Internet Backbone maps. Russ Haynal's ISP Page provides a comprehensive list of links to ISP backbone network maps. Also has details on Internet interconnection points.
- Visual Browsing in Web and Non-Web Databases, an excellent summary report ".. of projects, research, products and services that describe or apply information visualization technologies for enhancing use and access to Web databases". Compiled and maintained Gerry McKiernan and Peter J. Wasilko, Iowa State University, USA.
- The International Telecommunications Union.
- The World-Wide Web Consortium.
- Conferences :
  - The INET conference proceedings are available online again.
  - The proceedings to the ACM Computer-Human Interaction (CHI) conferences are online :
  - CHI95 Conference, "Mosaic of Creativity", 7-11th May 1995, Denver, USA.
  - CHI96 Conference, "Common Ground", 13-18th April 1996, Vancouver, Canada.

⊗ Done   ⊗ Internet

Commerce likes to know you (a lot)

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

E-mail Campaigns

BLUE MARTINI

- **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"**

19

(source: http://robotics.stanford.edu/users/ronnyk/etailKohavi.pdf)

Linking data to give combined portrait
According to Ron Kohavi director of data mining and personalization at Amazon.com

Demographic Overlays

BLUE MARTINI

- **Combine information about your customers from all possible sources**

- **A great source that is often overlooked is demographic overlays**

- **Very easy and cheap (8 cents/name) to get basic attributes like gender, income, profession, own/rent, car type, etc.**
  Note: not very reliable per person, but good for averages and segmentation

© Copyright 2002, Blue Martini Software. San Mateo California, USA

24

(source: http://robotics.stanford.edu/users/ronnyk/etailKohavi.pdf)

Power of maps

1. seeing patterns (eureka moment)

2. revealing process
- (proving theory to others)

3. classifying space and people. enables geographically specific action

Dr John Snow's 'Cholera' Map, 1854

Charles Booth's 'Inquiry into the Life and Labour of the People in London', 1880s

1898-99 Map
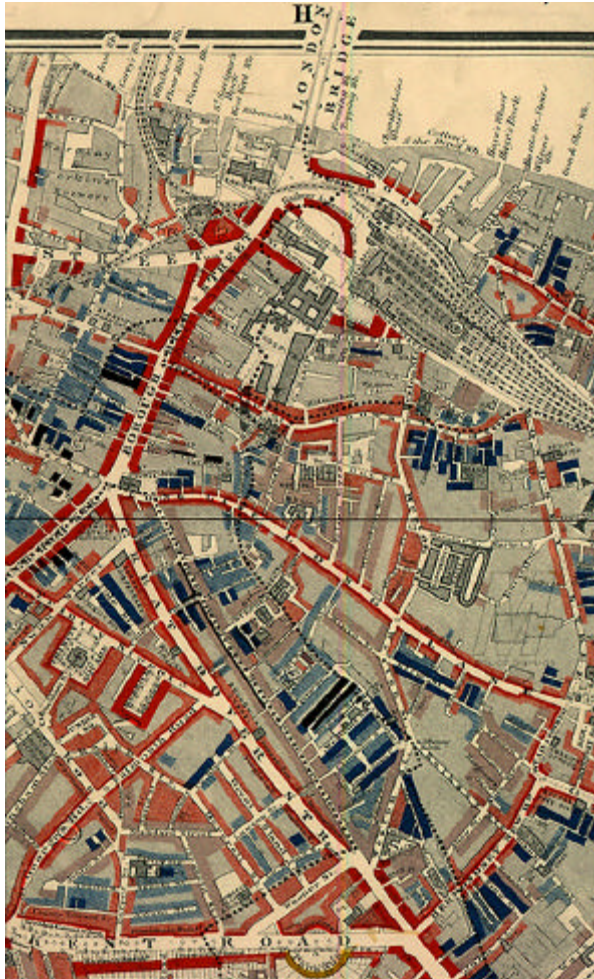
▲Reduce view area

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

Lowest class.Vicious, semi-criminal.

Very poor, casual. Chronic want.

Poor. 18s. to 21s. a week for a moderate family.

Mixed.Some comfortable, others poor.

Fairly comfortable.Good ordinary earnings.

Middle-class.Well-to-do.

Upper-middle and Upper classes.Wealthy.

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.

(Source: www.umich.edu/~risotto/)

MAPS FOR WEB SURVEILLANCE

Difference mapping makes

- visual data mining of social cyberspace
- opportunities to combine powerful capabilities of software and people, exploits best abilities of both
- fast but dumb calculation and record keeping versus slow but smart recognition and interpretation
- case studies
  1. hyperlink structures
  2. clickstream visualisation
  3. list participation / posting
  4. chat dynamics
  5. social networks in email
  6. activities in 3d virtual worlds
- not designed for surveillance in a negative sense, but raise privacy concerns in their application

1. Hyperlink structures

2. Web browsing patterns, clickstream visualisation

- data mining is automated extraction of hidden predictive information from large databases (e.g. web logs & transactions)

- 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

(a) Males

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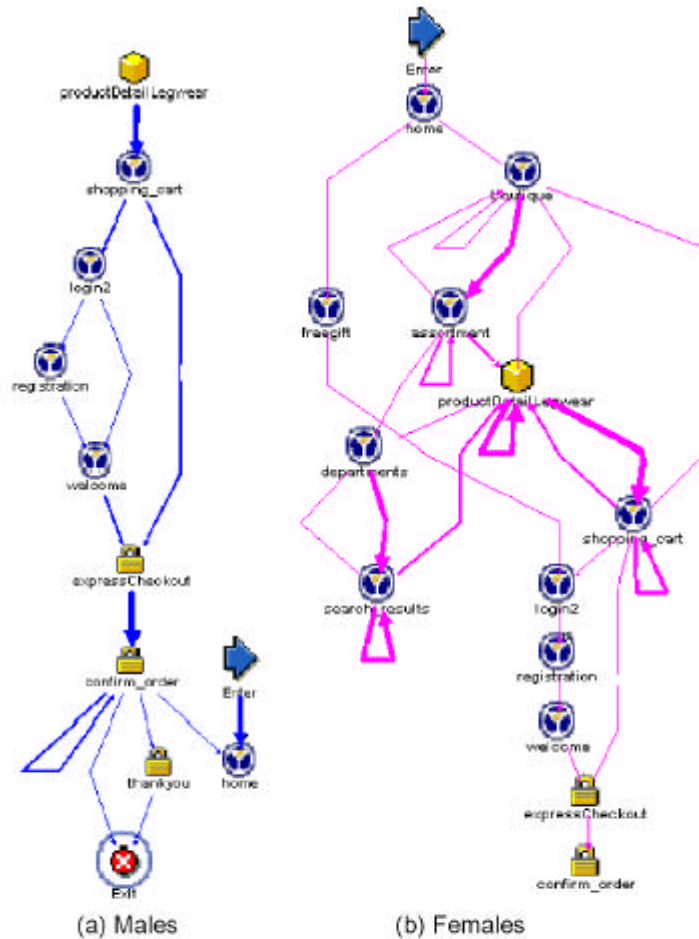
Figure 3. Gender Differences: Males tend to navigate in specific, direct patterns, whereas women's navigation patterns include much more browsing, utilizing much more of the site.

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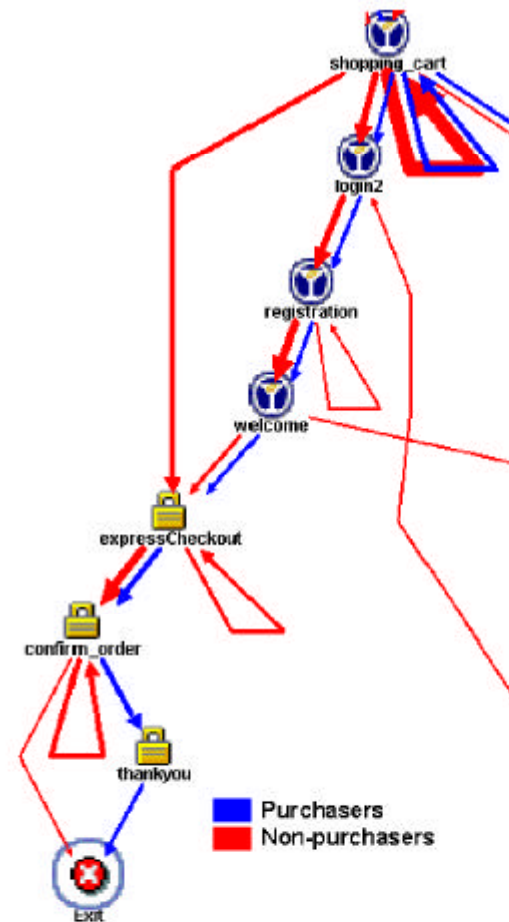
Figure 4. Checkout process. Purchasers take a direct route through the checkout process, whereas non-purchasers show a more haphazard route, including self-edges and early abandonment, possibly indicating a confusing checkout process.

(Source: Brainerd, J. and Becker B., 2001. "Case Study: E-Commerce Clickstream Visualization")

VISVIP
by John Cugini

(Source: http://zing.ncsl.nist.gov/WebTools/VisVIP/overview.html)
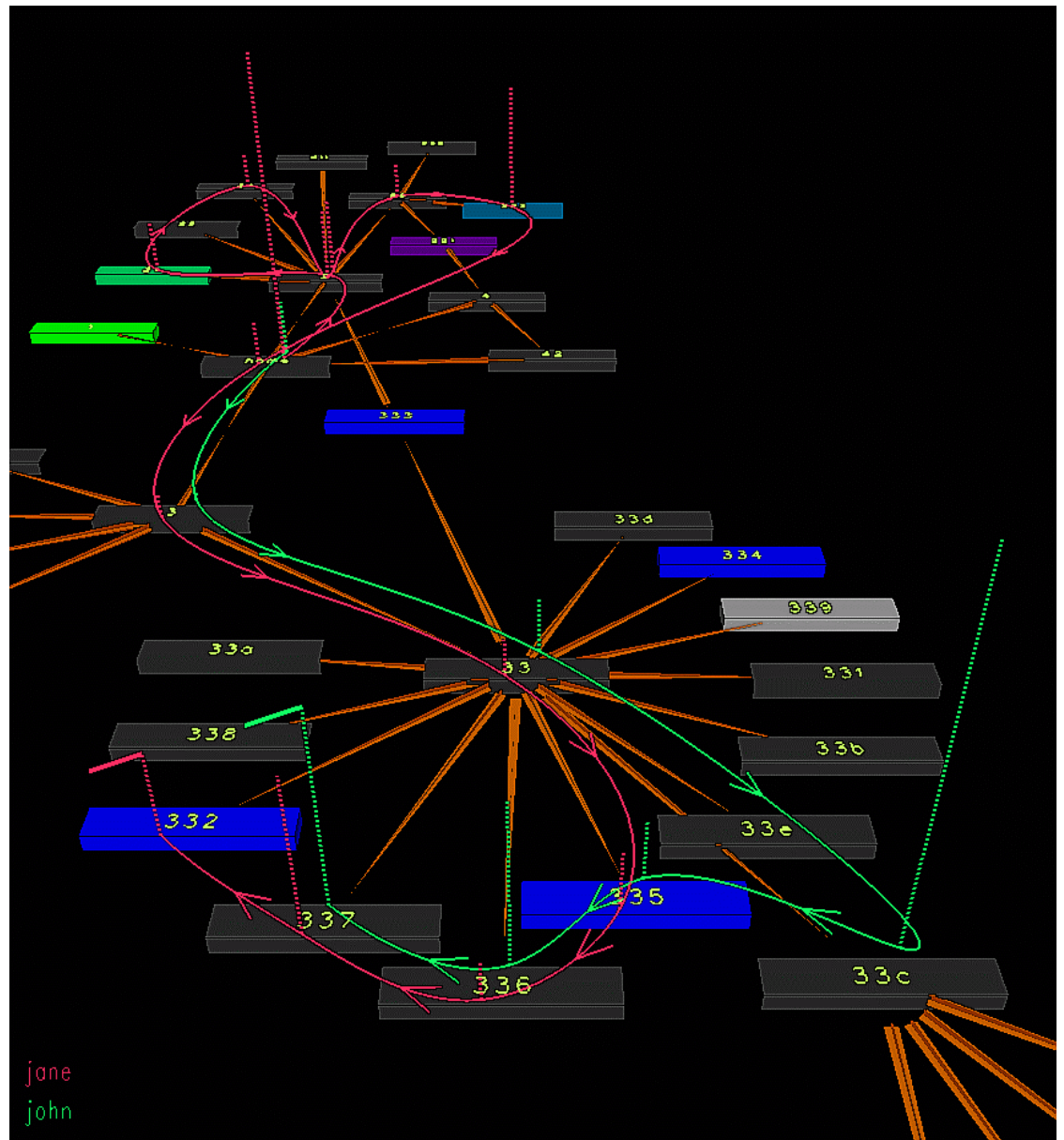
3. Community structure and individual participation

Mailing list affinities

VisualWho by
Judith Donath, MIT
Media Lab

(Source: http://smg.www.media.mit.edu/projects/VisualWho/)
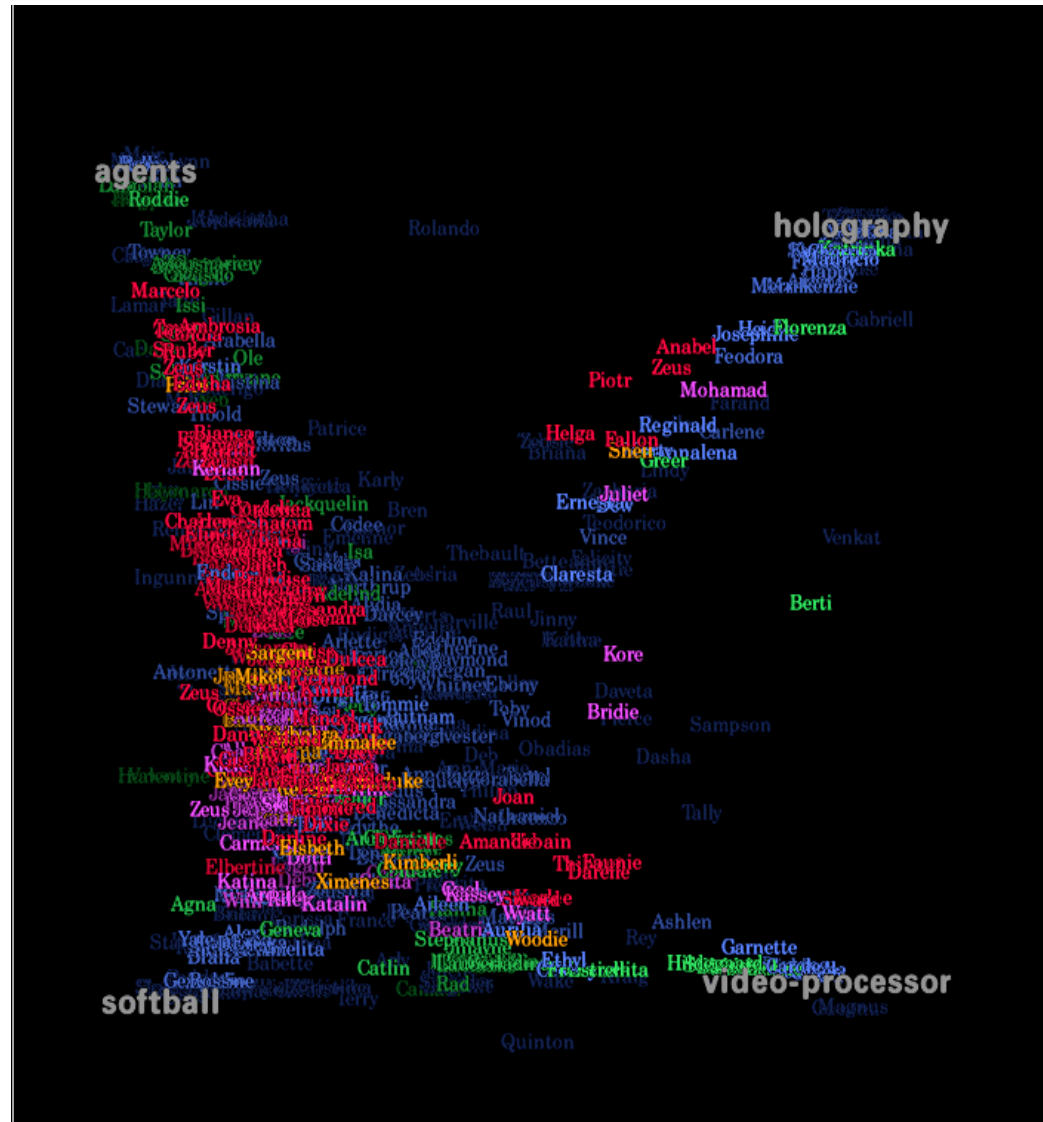
4 . Individual movement and activities in 3d virtual worlds
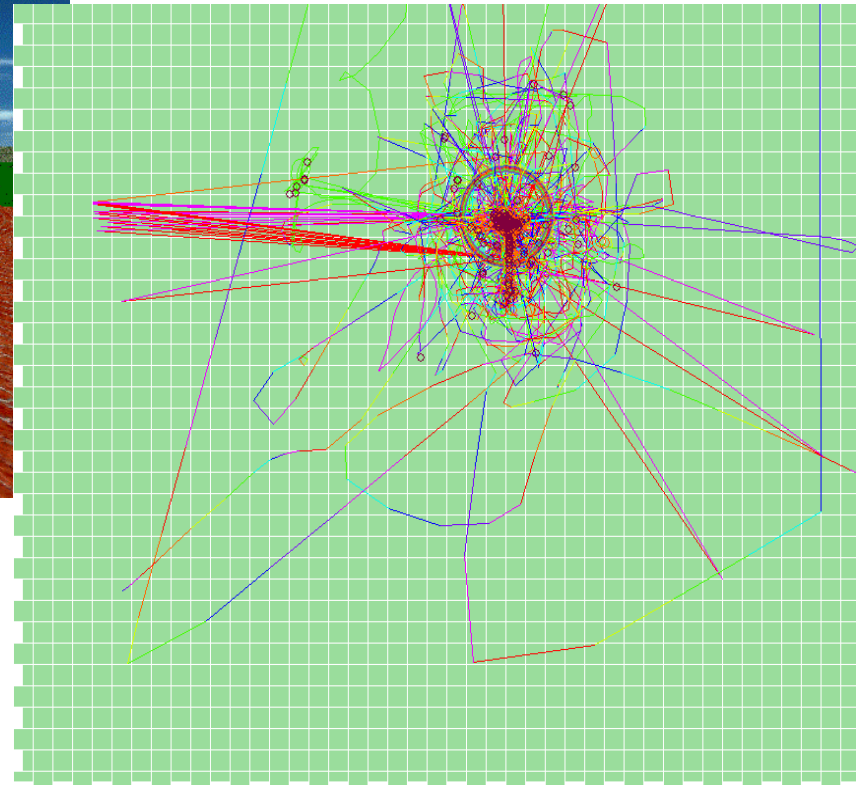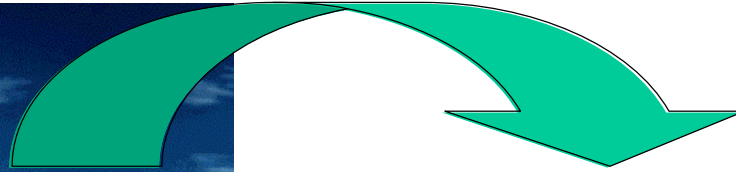
Interactive visualisation tool for analysing user trails in ActiveWorld space.
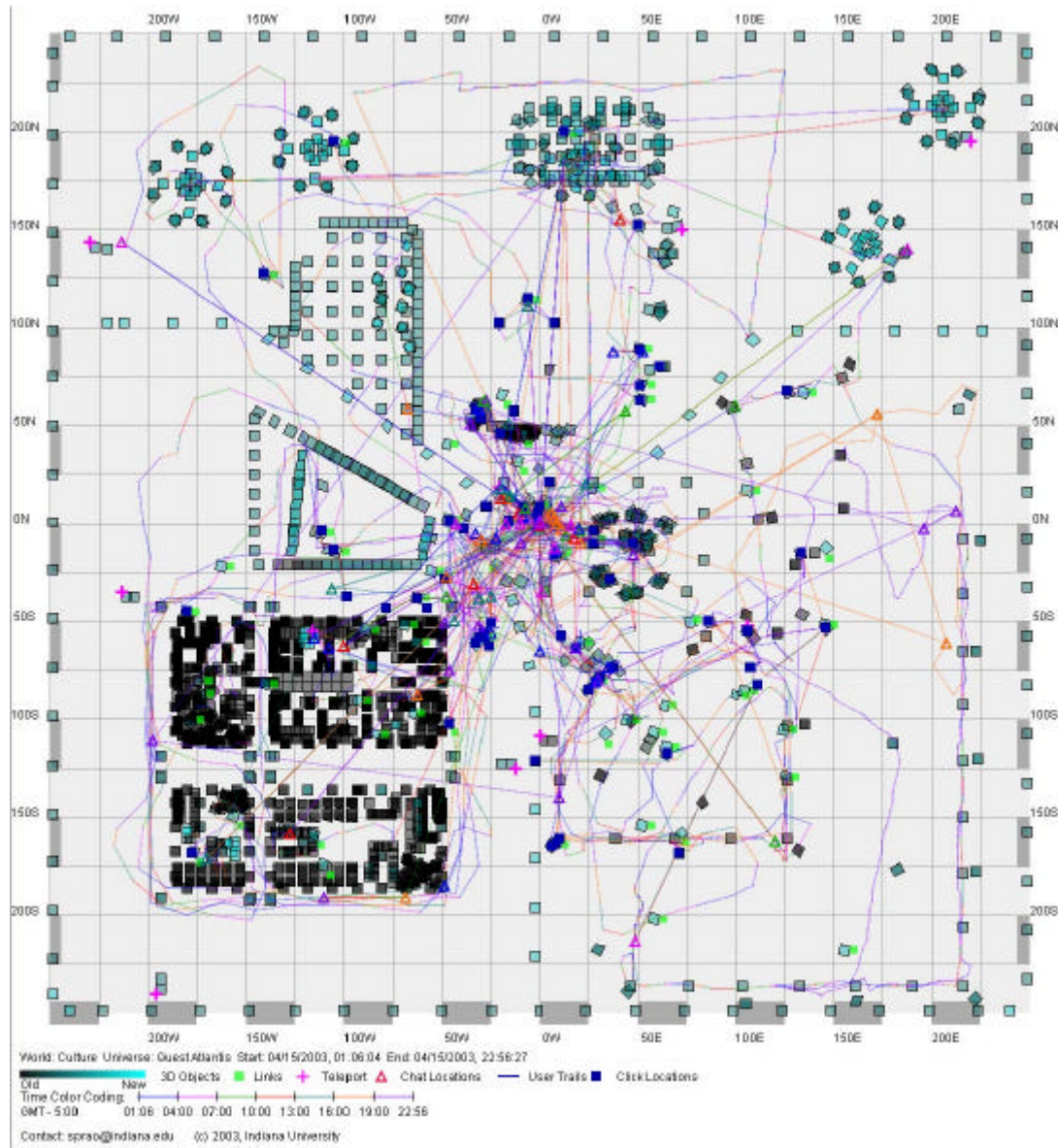
Katy Börner, Indiana University http://ella.slis.indiana.edu/~katy/

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

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

3D Objects ■ Links + Teleport △ Chat Locations — User Trails ■ Click Locations

Contact: sprao@indiana.edu   (c) 2003, Indiana University
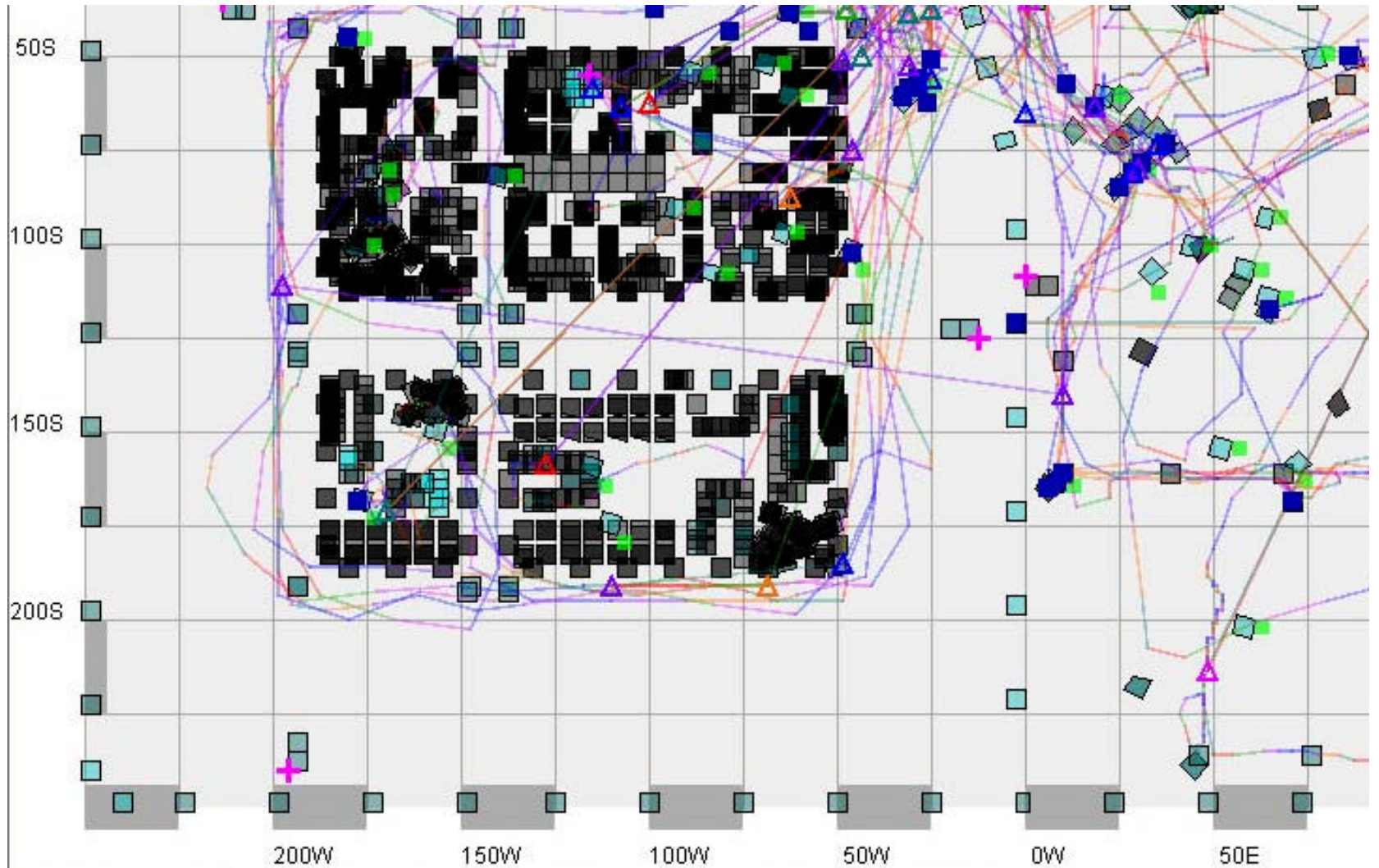
Beyond individual privacy rights

- surveillance grants the watcher greater power of the watched

- negative implications of censorship and discrimination (informational, financial and material)

- feedback effect - narrowing of choice

- enabling better blocking and censorship. chilling of cyberspace by DMCA and post 9-11 security paranoia

- decline in ability to read anonymously. attempted anonymity will be seen as marker of suspicion and risk

- dangers of self censorship

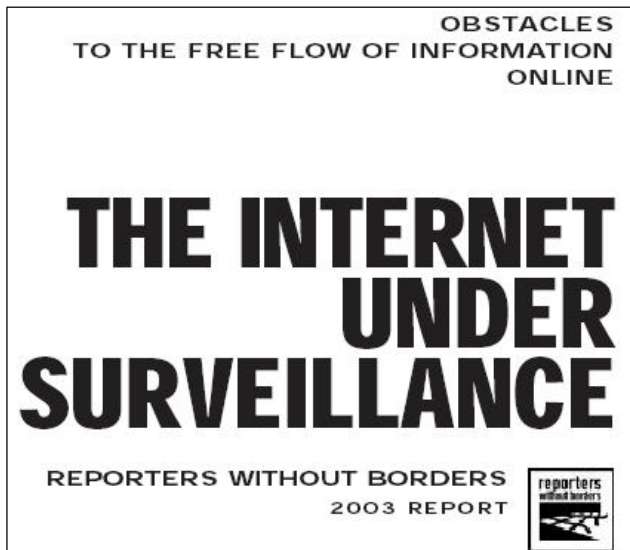# We don't like your browser, you can't come in

"Exclusively for Everyone (except mozilla users)"

Ethics of researchers and their tools

- maps of social cyberspaces are 'responsible artefacts', that do not destroy what they seek to represent or enhance.
- Marc Smith notes, '[t]he bright light of social science research can create an unpleasant glare for participants drawn to a dimly lit online space.' (from Invisible crowds in cyberspace: Mapping the social structure of the Usenet)

- Unrelated activities tied together into unified dossiers, sum of the parts of more revealing

- easily drawn into governmental security systems, commercial consumption profiling systems
- moving from discretionary to mandatory

- 'software sorted society' based on their activity patterns

- questions ?? I would welcome feedback to m.dodge@ucl.ac.uk
- the slides of this presentation are available at http://www.casa.ucl.ac.uk/martin/paris04.pdf
- Terra Australis Incognita map courtesy of National Library of Australia http://nla.gov.au/nla.map-nk2456-13
- further reading:

OBSTACLES
TO THE FREE FLOW OF INFORMATION
ONLINE

THE INTERNET UNDER SURVEILLANCE

REPORTERS WITHOUT BORDERS
2003 REPORT

reporters
without borders

silenced

an international report on censorship and control of the Internet

By Privacy International and the GreenNet Educational Trust

www.rsf.org   www.privacyinternational.org/survey/censorship/