Centre for Advanced Spatial Analysis

Census and Related Information by Web Visualization

Attracting Millions of New Users?

Michael Batty
m.batty@ucl.ac.uk
Steven Gray, Richard Milton & Ollie O'Brien
steven.gray@ucl.ac.uk richard.milton@ucl.ac.uk o.o'brien@ucl.ac.uk

http://blogs.casa.ucl.ac.uk/

CGG Seminar Working Party
Generating Value from the 2011 Census
19 September, 2012

Centre for Advanced Spatial Analysis

Outline

- The Context: A Dramatic Change in Media for Display and Dissemination
- Web 2+ and Online Mapping
- Integrating Diverse Data: Adding Value
- Generating New and Complementary Data: Crowd-Sourcing
- Visual Analytics
- Census Analytics
- Next Steps: What Can We Expect?

The Context: A Dramatic Change in Media for Display and Dissemination

From Ollie O'Brien's talk to this group last year, I retrieved the following map of visualising population density in the 1841 UK Population Census.

The map was online at ….

http://pressandpolicy.bl.uk/Resource-Library/Augustus-Petermann-s-population-density-map-1841-credit-British-Library-Board-34a.aspx

And I pulled it up to find that I could download it from the British Library web site. This was a real surprise as I hadn't seen the map before and it illustrates that the web not only allows us to discover the present and the future but the past. I was so taken with the map I downloaded it to my iPad and then my wife took a picture of it and me with my iPhone, I emailed it to myself and then transferred it here to my MacAir

Centre for Advanced Spatial Analysis UCL

It's great example of the power of the web to disseminate and connect through many devices and media

Centre for Advanced Spatial Analysis UCL

To an extent, much if not most of what I will talk about relates to how new technologies are changing the way we are able to display and disseminate.

Web 2 is essentially a medium in which we can interact with data in an online context, consuming and producing new data

Enormous strides are being made in this new world but we do need to exercise a degree of caution. Our science is being changed for sure but it is not clear that it is getting better. Our analytics as we now call them, are not moving as fast as our ability to visualise.

Centre for Advanced Spatial Analysis

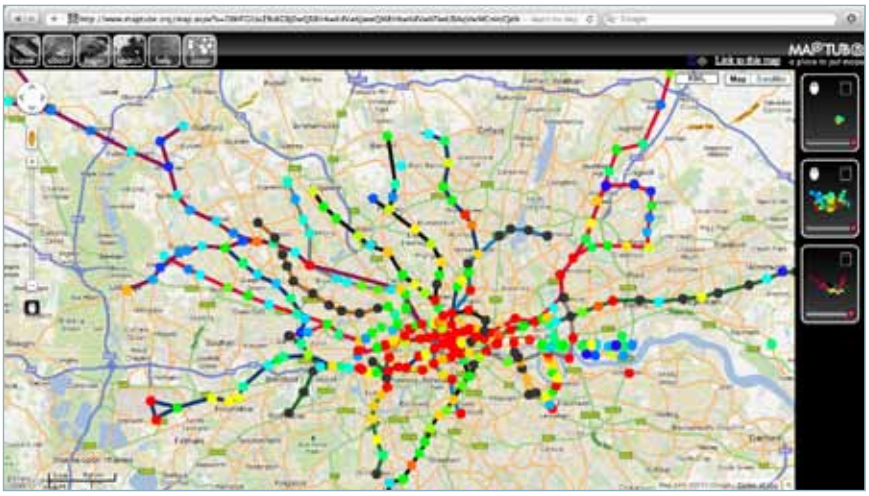
I am going to begin with looking at the sorts of maps we can now visualise online and there are many of these sites where one can do this and interact to produce simple analysis and new data

We usually do it on the desktop, but there are various applications to smaller devices such as tablets and phones and doubtless when we get digital paper, we will have yet another medium to enable us to create new ways of interaction.

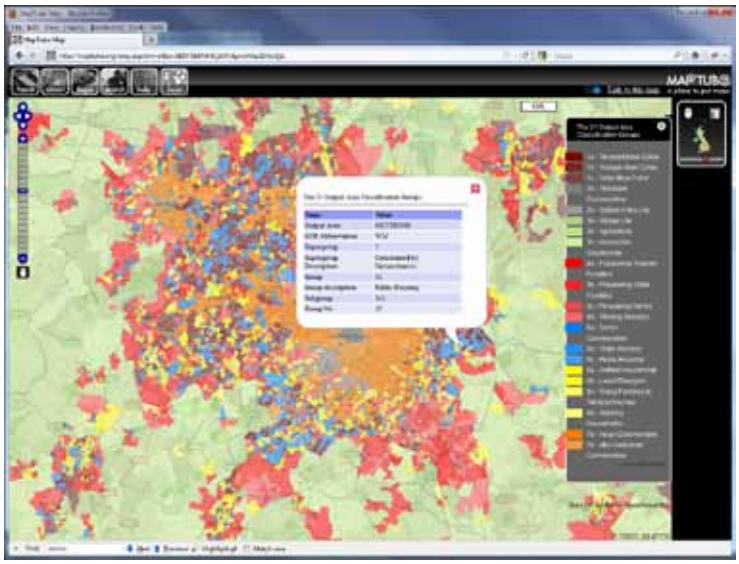
I am going to start by showing you our entry into this world some 5 or 6 years ago when we created our site called Maptube: www.maptube.org

Centre for Advanced Spatial Analysis

Web 2 and Online Mapping

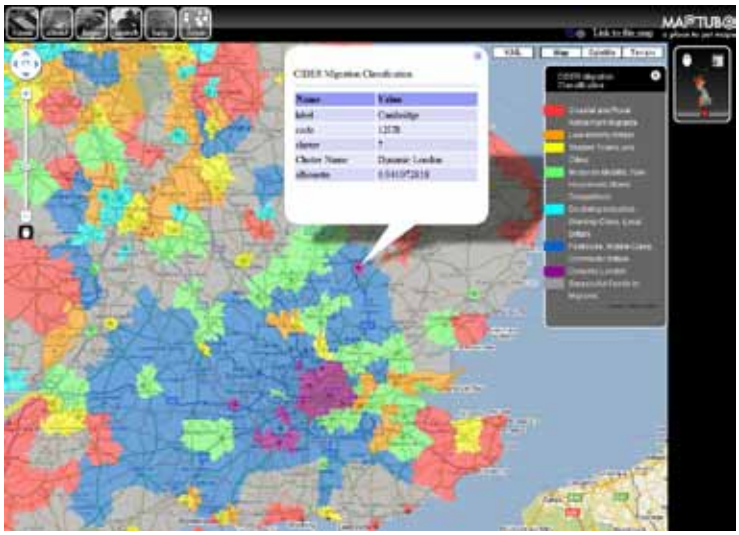


Centre for Advanced Spatial Analysis



Map data CC-By-SA OpenStreetMap, Census data Copyright ONS. Shown using MapTube by Richard Milton

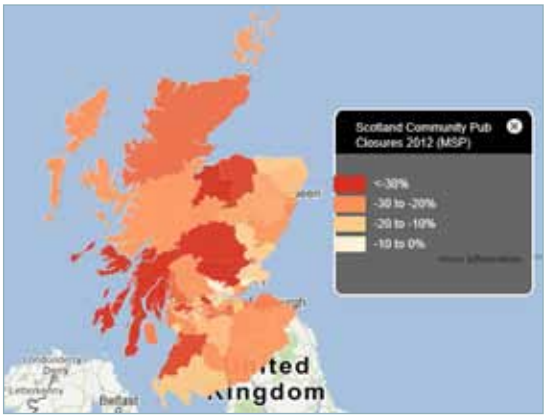
Centre for Advanced Spatial Analysis



Produced by Adam Dennett (CASA) – Census data Copyright ONS –
see http://www.adamdennett.co.uk/ and http://www.maptube.org/

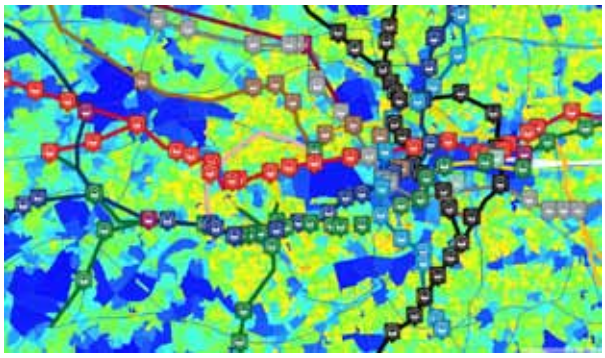
Centre for Advanced Spatial Analysis

May 2010 General Election Result



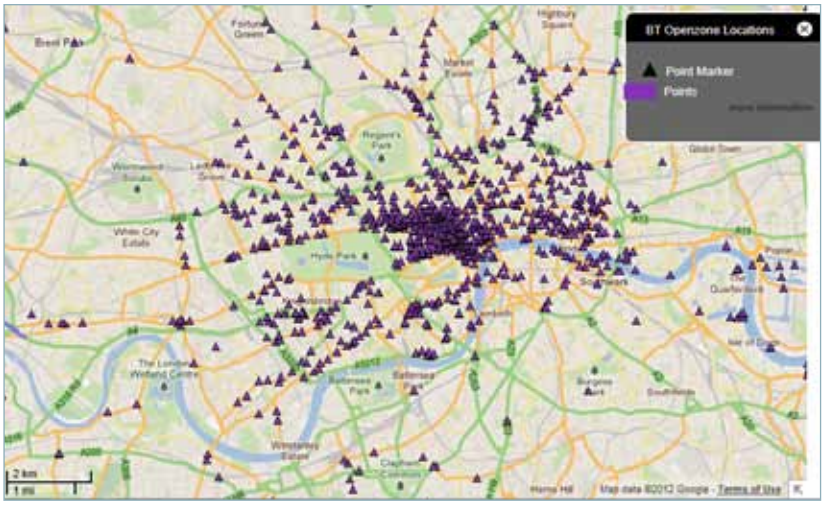
Scotland Community Pub Closures, source: CGA Strategy

Centre for Advanced Spatial Analysis



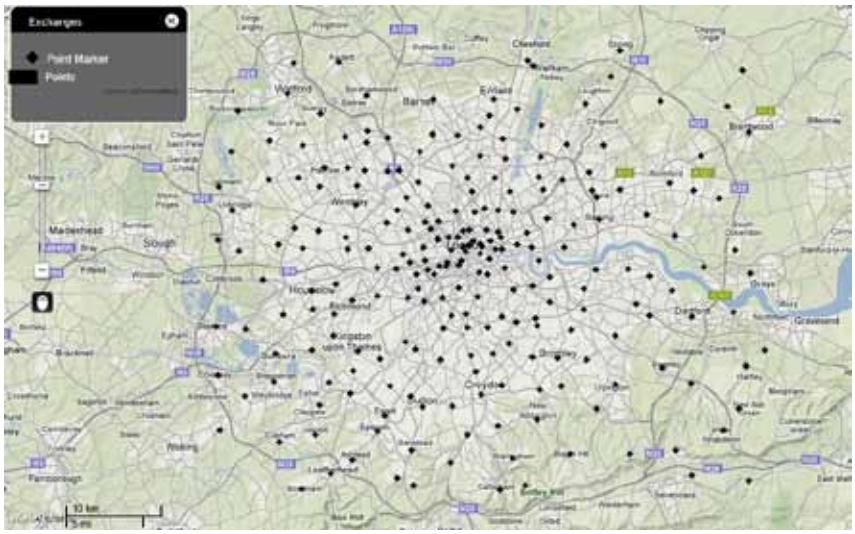
Adding Value – Comparing Data Sets
Population density (2001 Census) with tube lines and real-time tube positions

Centre for Advanced Spatial Analysis



Information Infrastructure: BT Openzone Hotspots, March 2012

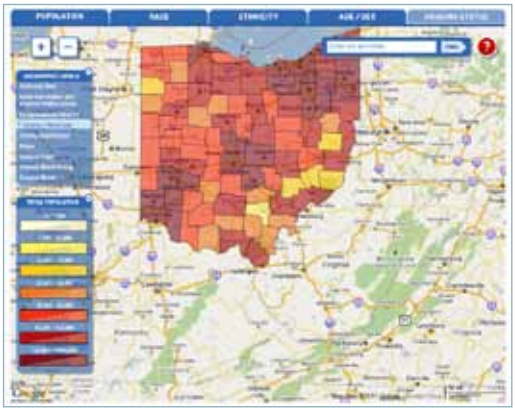
Centre for Advanced Spatial Analysis



Information Infrastructure: Telephone Exchanges

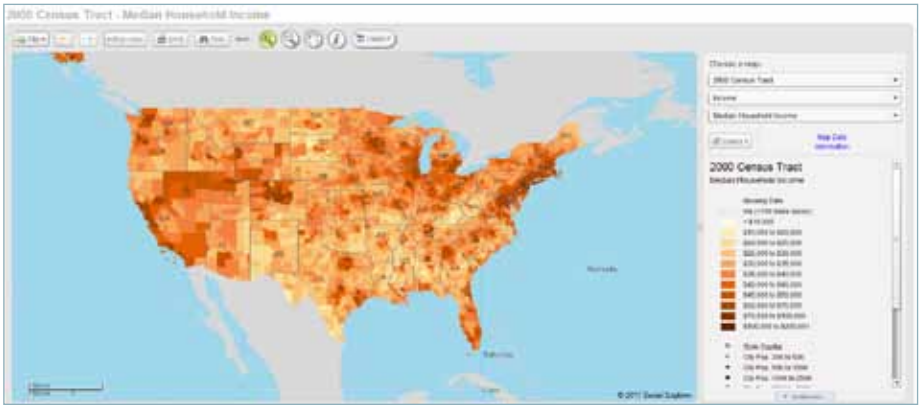
There are a whole variety of web sites now enabling you to explore maps based on various mashups. I will show some from the 2010 US Census stimulated by its release.

http://www.zerohedge.com/article/interactive-visualization-2010-census-results



http://2010.census.gov/2010census/popmap/index.php

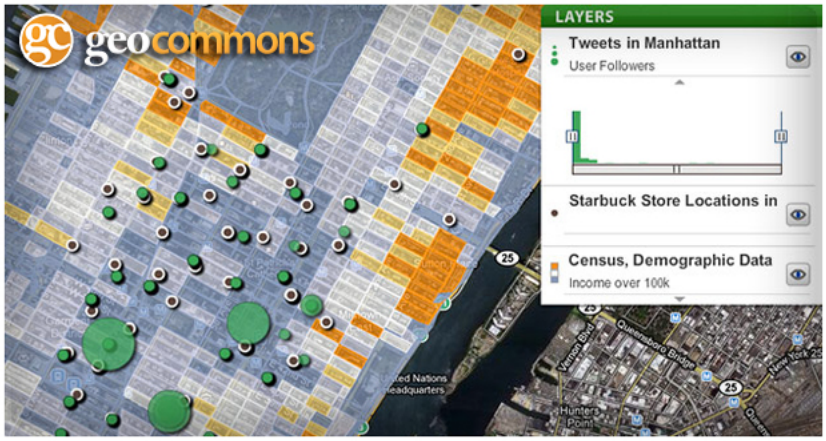
Centre for Advanced Spatial Analysis



Social Explorer Free edition

http://www.socialexplorer.com/pub/maps/map3.aspx?g=0&mapi=SE0012

Centre for Advanced Spatial Analysis



http://multimedia.journalism.berkeley.edu/tutorials/build-interactive-census-map-geocommons/

Centre for Advanced Spatial Analysis

Our Census Profiler

Image-based
- Faster on the web browser
- Not delivering restricted data to the web browser

Open Source software
- Leverage the powerful OpenLayers mapping software
  - More powerful than Google Maps equivalent
- An active development community
- Full access to the underlying code – greater flexibility

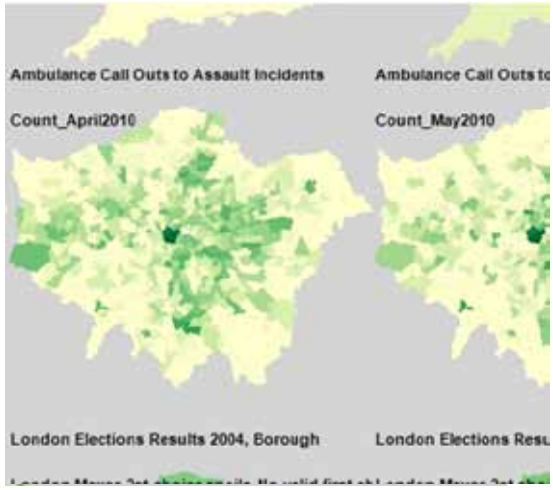
"Slippy" map
- Intuitive
- Encourages exploration

Map data CC-By-SA OpenStreetMap, Aerial imagery Copyright Google, Census data Copyright ONS

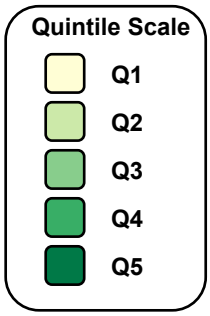
Centre for Advanced Spatial Analysis

Map data CC-By-SA OpenStreetMap, Aerial imagery Copyright Google, Census data Copyright ONS

Centre for Advanced Spatial Analysis



102 datasets
1706 variables (~50%)

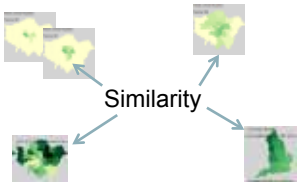


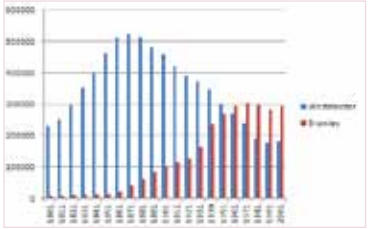
london.gov.uk

Mining a Datastore: Ordering and Visualising Data

Centre for Advanced Spatial Analysis

London Datastore Historic Census Population: 1801, 1841 and 1939





Mining a Datastore: Historic Populations for London

Centre for Advanced Spatial Analysis

UCL

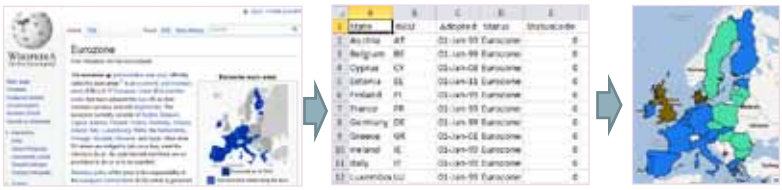
Integrating Diverse Data: Adding Value

Using our MapTube resource developed under DSR Genesis programme, we can now extract many open data

Centre for Advanced Spatial Analysis

UCL

Centre for Advanced Spatial Analysis



Algorithm

Polygons

1. For the first N rows (N=10,000) of every column, use a RegEx test to rule out any columns that can't possibly match
2. For the first N rows (N=1,000) of the remaining columns, lookup key text in geocode database containing tuples of (key,dataset name) for every geography
3. Assign probability to column (prob, dataset) tuple based on number of matched rows

Points

1. Compute statistics on columns for: Min, Max, IsNumeric, IsProgression and Column Name Weight
2. Find X and Y Columns based on IsNumeric&!IsProgression
3. Choose CRS based on Min and Max

We are developing many new versions of MapTube that enable us to extract data from web sources

Centre for Advanced Spatial Analysis

Generating New and Complementary Data: Crowd-Sourcing

We will show some early work dating to the beginning of the credit crunch. But the classic example is Open Street Map, that we are proud to say has its origins in UCL, if not in CASA where Steve Coast was an intern – terrible word – in the late 1990s

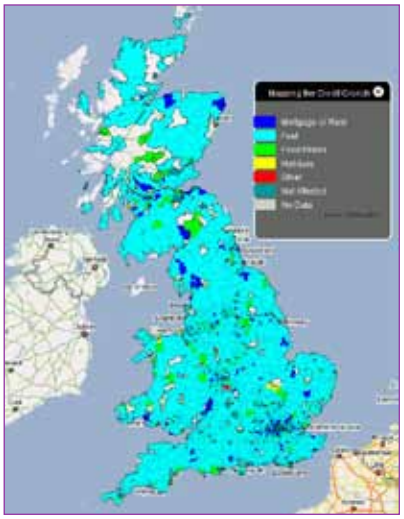
To generate data through eliciting responses to questions or ideas via the web, one needs a good broadcast medium and this is essential – our entry into this domain was through the BBC who came to use, having seen MapTube and asked us to help them elicit responses to how people felt about the credit crunch

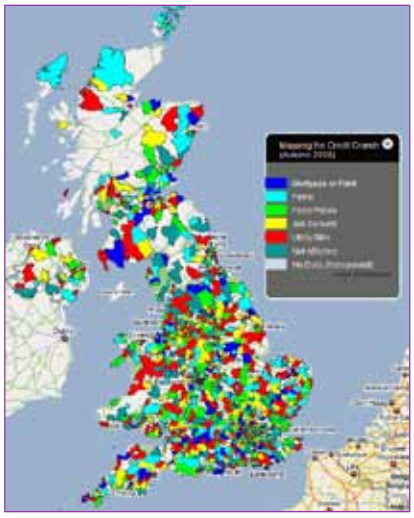
Centre for Advanced Spatial Analysis

23,475 responses
April, May, June 2008

A new credit crunch survey started in October and currently has 3,802 responses.

Centre for Advanced Spatial Analysis

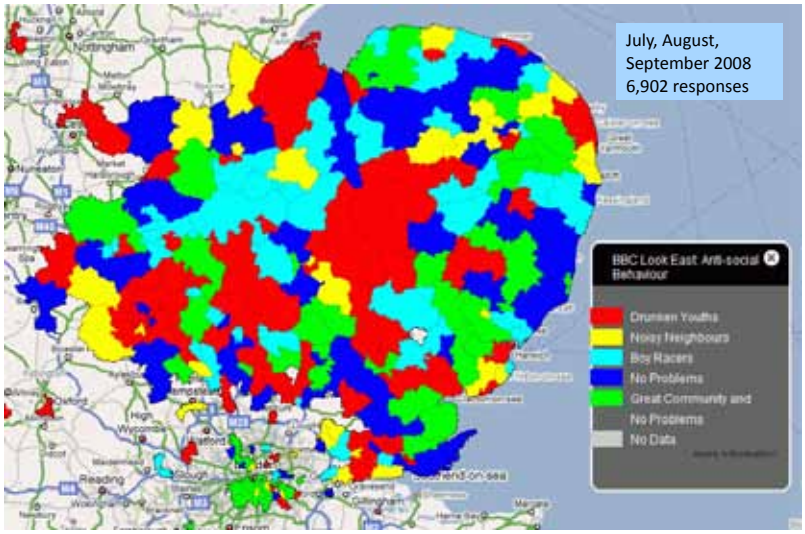




http://www.maptube.org/creditcrunch/

BBC Look East: Anti-Social Behaviour

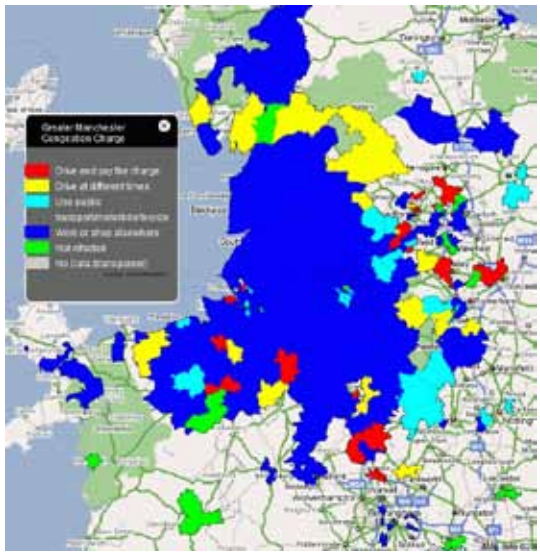
Centre for Advanced Spatial Analysis



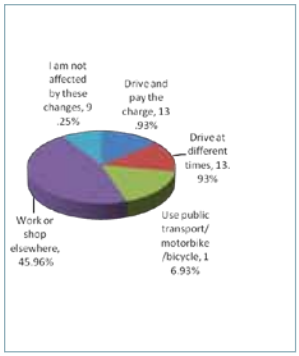
http://www.maptube.org/lookeast

Manchester Congestion Charge

Centre for Advanced Spatial Analysis

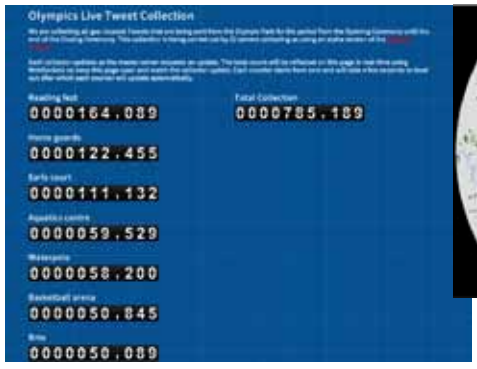


15,902 responses
October to December 2008



Data, Sensing, Capture, Extraction:

Crowd-Sourcing: Survey Mapper let's you create a survey and mount it on the web; this is part of the BigDataToolkit



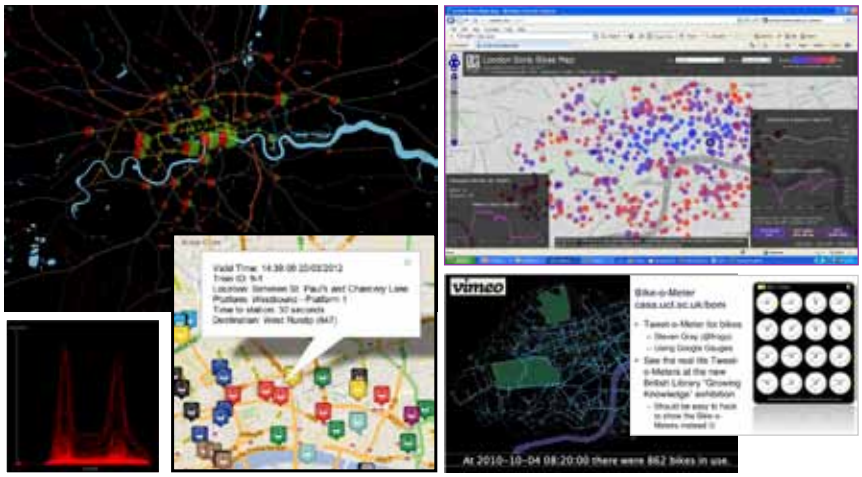
Visualising London's Tweets
3 Months – Jan to March 2011

Collecting social media using open APIs such as Twitter data; extracting meaning from such data such as spatial locations. The collection requires large storage (multi-server) capacity for routine analysis

Centre for Advanced Spatial Analysis

UCL

Visual Analytics

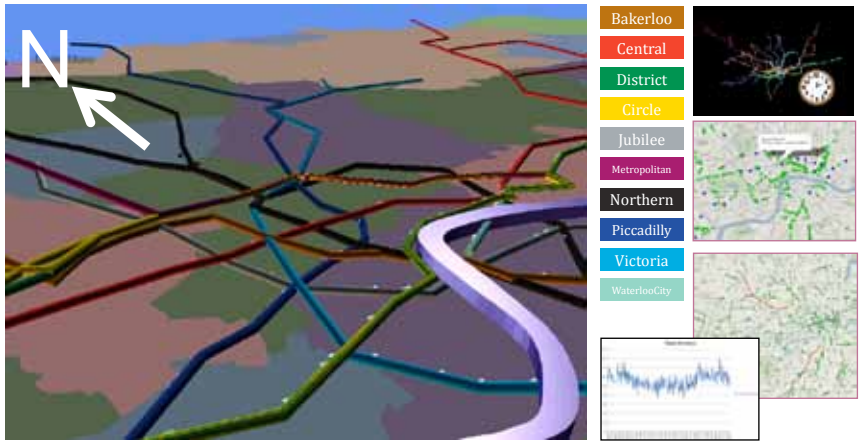


Collecting Oyster card, public transport flows, bike use etc from very large archives of open data (~ a billion records)

Centre for Advanced Spatial Analysis

UCL

Visual Analytics:



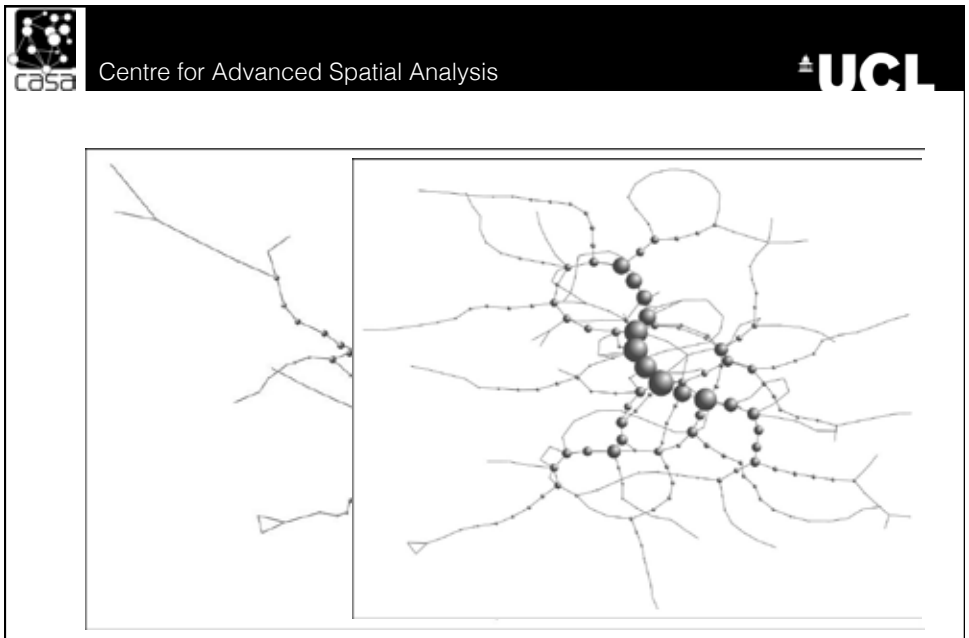
Visualising flows on networks from sampling and recording real-time movements

From Trackernet 8am to 8pm, 16 April 2012

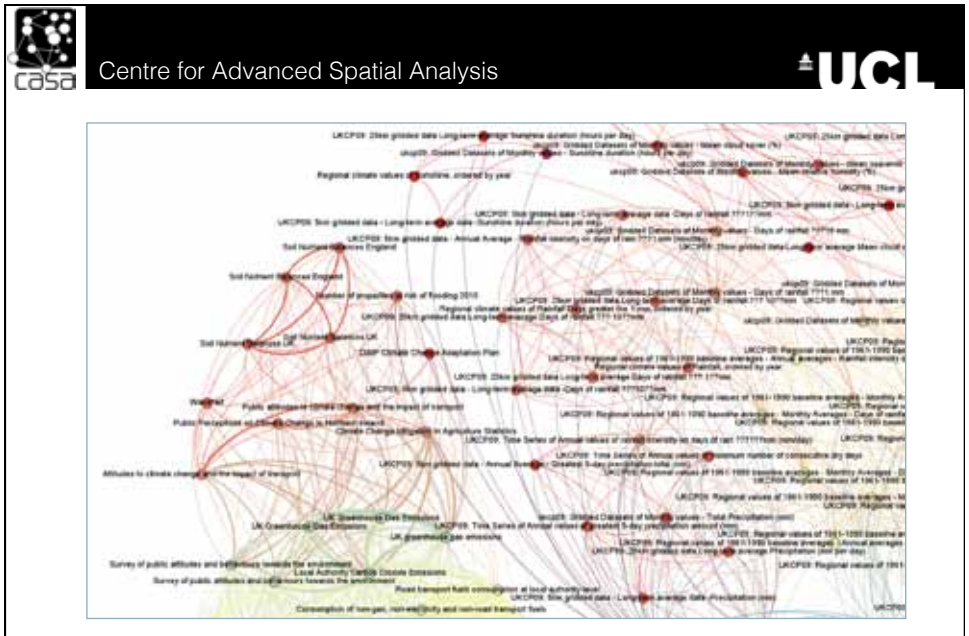
Centre for Advanced Spatial Analysis UCL

Thursday 21 June 2012 13:18 Friday 22 June 2012 09:00

Two images of bus positions on consecutive days showing the impact of the bus strike on East and West London.

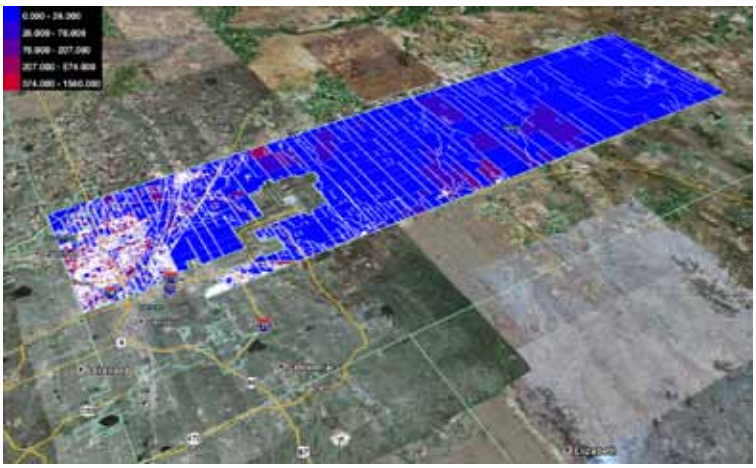


Visualising network structure: accessibility centrality as a planar graph, using force directed graphs



Visualising network structure: taking data on the web and exploring its structure: climate data

Census Analytics



http://www.gcensus.com/

Centre for Advanced Spatial Analysis

http://thistract.com/

Built by Michal Migurski of Stamen Design

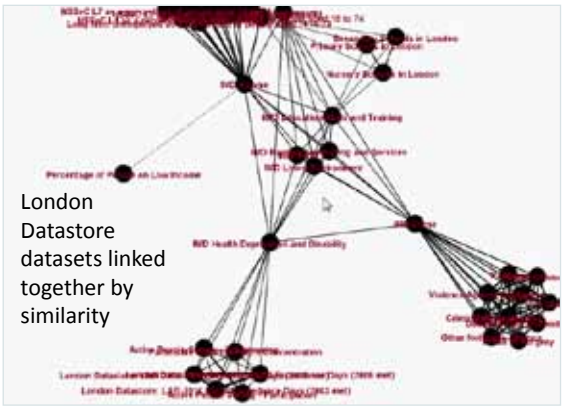
Uses geolocation

On "GitHub"

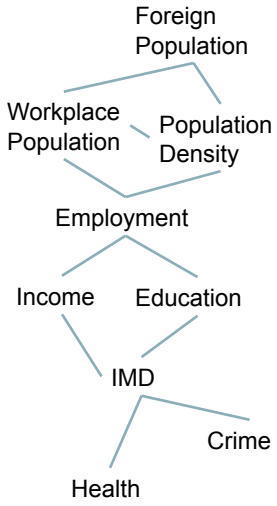
- Should be possible to implement it for the 2001 (and 2011) UK census data
- "This OA"?

Centre for Advanced Spatial Analysis UCL

Keyword Graphs: How are datasets linked together?



London Datastore datasets linked together by similarity



Centre for Advanced Spatial Analysis

http://www.gapminder.org/

Google Data Explorer

http://www.google.com/publicdata/home

Centre for Advanced Spatial Analysis

Textal: Visualising textual data from Books, Social Media, etc. Drill down into the Analytics

Census Profiler + Google Earth

Aerial imagery Copyright Google, Boundary data Crown Copyright, Census data Copyright ONS

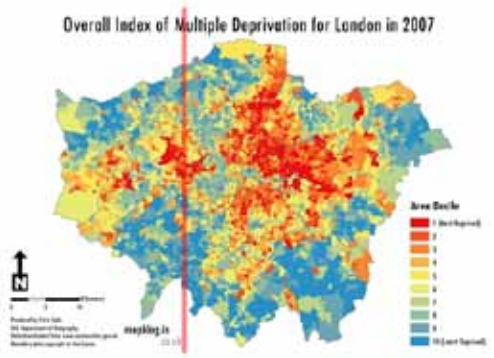
Centre for Advanced Spatial Analysis

Aerial imagery Copyright Google, Boundary data Crown Copyright, Census data Copyright ONS

Centre for Advanced Spatial Analysis

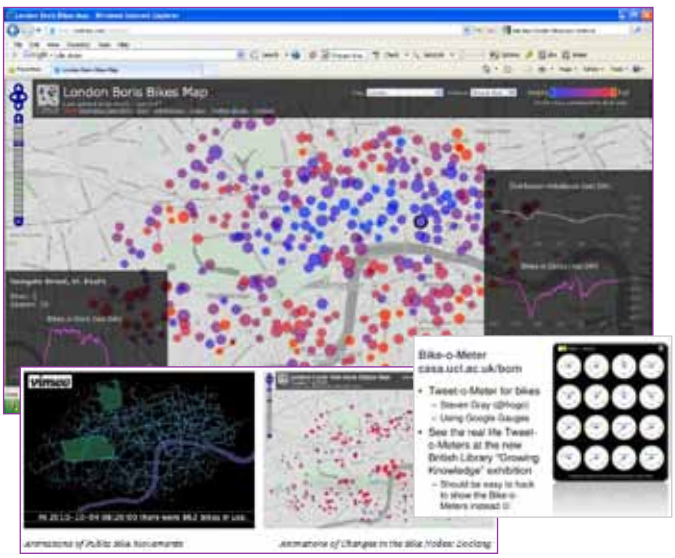
Next Steps: What Can We Expect?

Choropleths to show specific value or % differences
Heatmaps to show general patterns of differences
Fade between years
Swipe maps



Swipe map was produced by Chris Gale (UCL Geography)

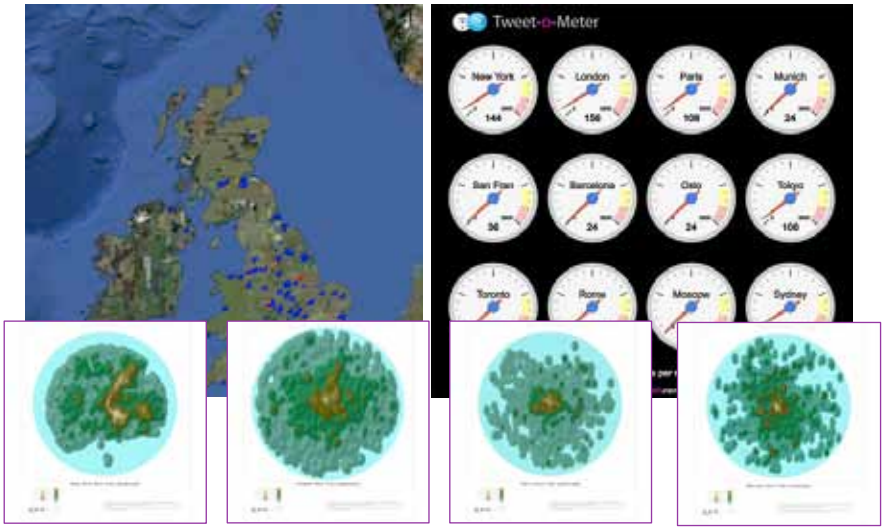
Centre for Advanced Spatial Analysis



Centre for Advanced Spatial Analysis

UCL

Spatial Analysis of Urban Activity using Twitter data

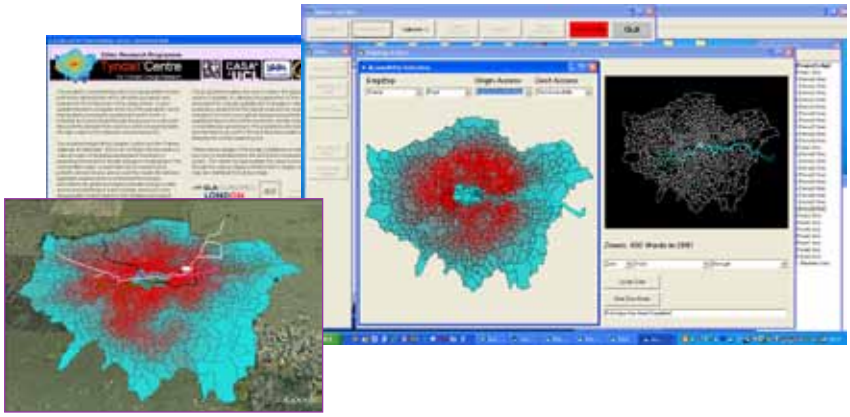


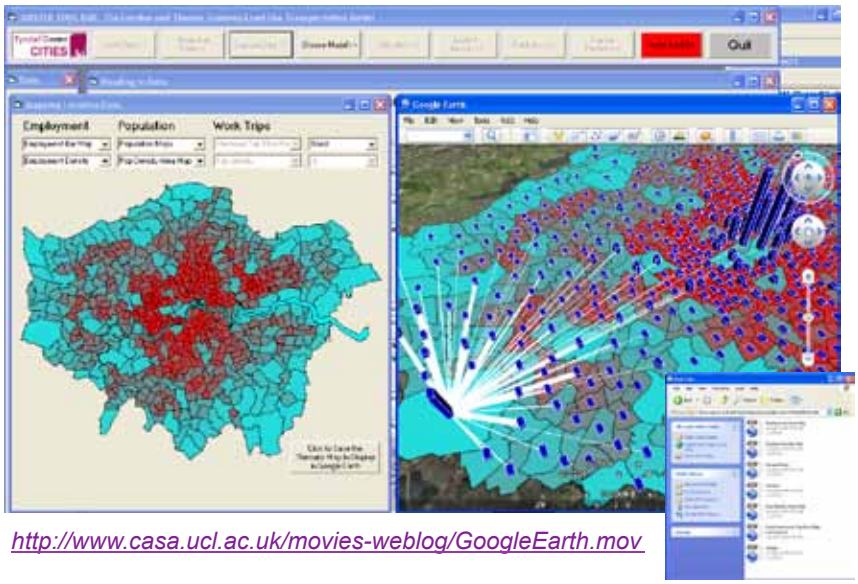
Centre for Advanced Spatial Analysis

UCL

Simulation: Modelling Land Use Transportation, Energy, etc

Our core expertise in CASA is also in land use transportation modelling and we have several such models for the London region:





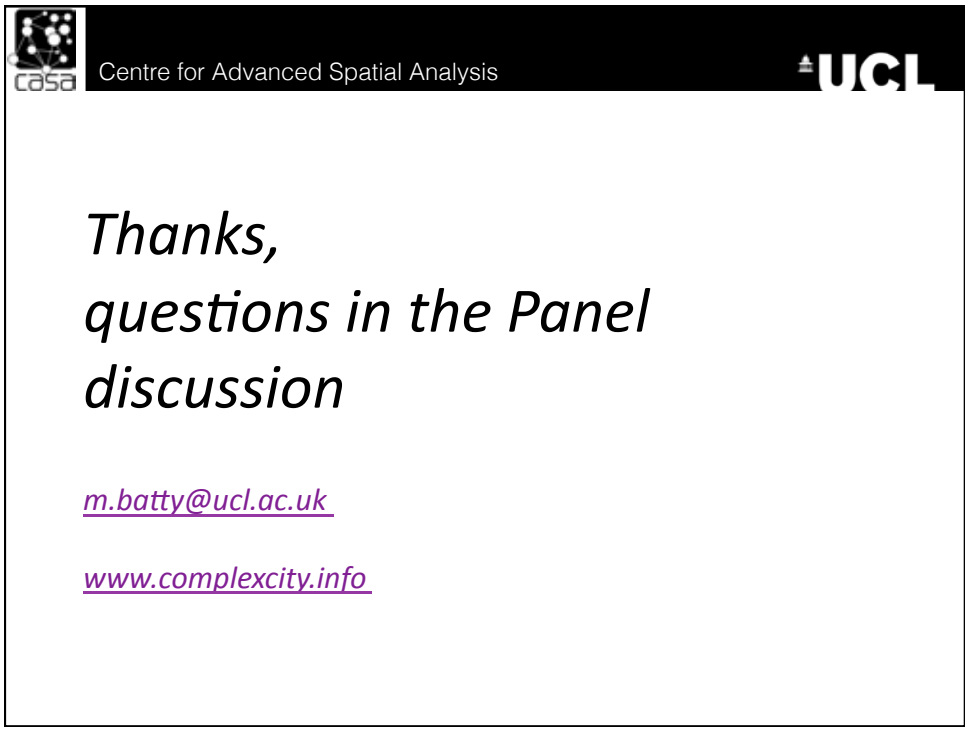
http://www.casa.ucl.ac.uk/movies-weblog/GoogleEarth.mov

Centre for Advanced Spatial Analysis UCL

Those born after 1993, have only known life with the web

"A generation whose first port of call for knowledge is the internet through Google's search engine, as opposed to books, libraries or traditional (off-line) information sources" (CIBER, 2008)

CIBER 2008: Information Behaviour of the Researcher of the Future.
http://pressandpolicy.bl.uk/Press-Releases/Pioneering-research-shows-Google-Generation-is-a-myth-32b.aspx


Centre for Advanced Spatial Analysis
UCL
Thanks,
questions in the Panel discussion
m.batty@ucl.ac.uk
www.complexcity.info