Big Data, Mobility, and Smart Cities

Michael Batty

m.batty@ucl.ac.uk
@jmichaelbatty
http://www.complexCity.info/
http://www.spatialcomplexity.info/

Saturday, 22 June 2019

The lecture is in three parts

- Smart Cities: Definitions, History, Big Data
- **Big Data, Mobility, and Smart Cities**
- Ideas About How we Use Computers in Developing Understanding and Planning Cities: Models

Big Data, Mobility, and Smart Cities

- Mobility, Transit, & Real-Time Streaming: The Oyster Card Data Set
- Learning about Mobility from the Data
- Variabilities – Heterogeneity and Travel Profiles
- Disruptions – Signal Failures, Stalled Trains
- Variable Locational Dynamics of Demand
- Related Real -Time Data: Bikes, Social Media
- What Can We Learn: The Limits to Big Data

Real-Time Streaming: The Oyster Card Data Set

- Tap at **start** and **end** of train journeys
- Tap at **start only** on buses
- Accepted at 695 Underground and rail stations, and on thousands of buses
- **Many Variants of the Data Sets**
- **991 million** Oyster Card taps over Summer 2012 – this is big data
- Quality of Data
- What Can We Use It For
- Missing Data and Noise





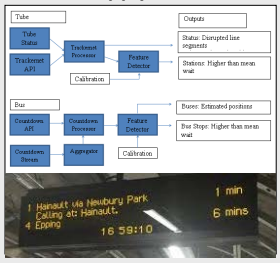


Real-Time Streaming: Oyster Card & Trackernet Data

Data Capture:

Demand by Travellers ***Supply of Trains***



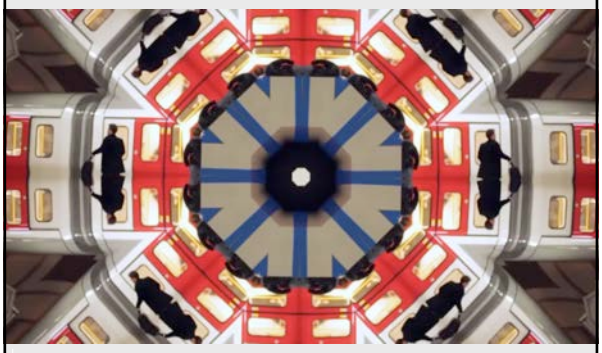


How do we match demand & supply without passenger tracking?

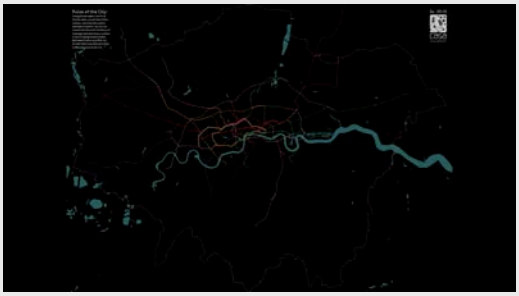


Tube, Overground and National Rail Networks in London where Oyster cards can be used

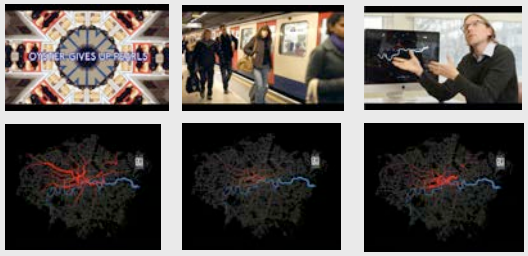




And how can we make sense of this



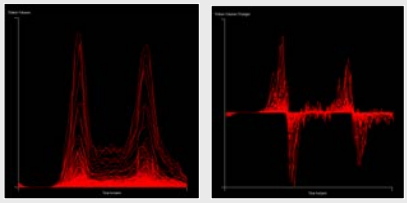
http://www.simulacra.info/

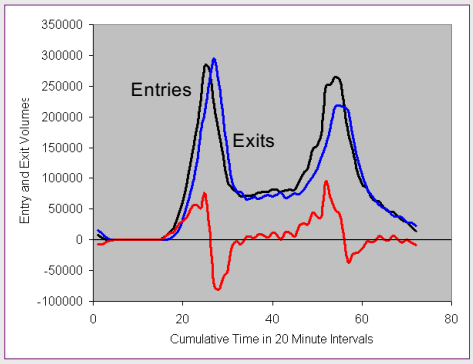


This of course was the thing that Lt Henry Harness did in Dublin in 1837 and what Minard et al. did a little later. In our LUTI models, this is an enormous problem as the scale of this assignment to networks is different

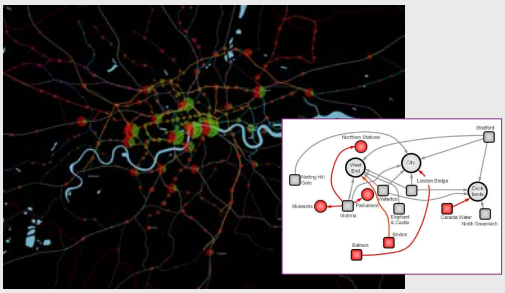
Variabilities – Heterogeneity and Travel Profiles

First we will look at some of the data and how it varies in terms of the diurnal flows usually morning and evening peaks, with a small blip (peak) around 10pm at night



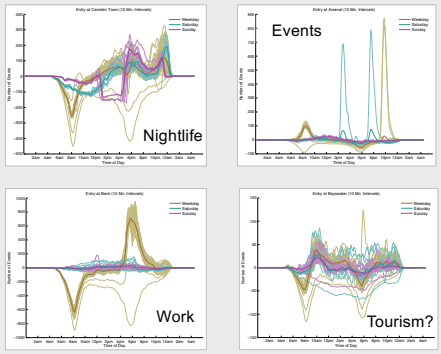


Oyster Card Data – interpreting urban structure, multitrips, etc.



Roth C., Kang S. M., Batty, M., and Barthelemy, M. (2011) Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. **PLoS ONE 6(1):** e15923. doi:10.1371/journal.pone.0015923

Particular Events: Weekdays, Saturdays and Sundays



Comparing Variability for different time intervals for Three World Cities: London, Beijing and Singapore

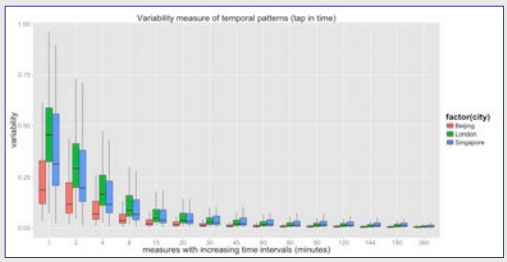
Table 1. Summary statistics of one-week of smart-card data (metro trips only)

	London	Singapore	Beijing
Monday	3,457,234	2,208,173	4,377,500
Tuesday	3,621,983	2,250,597	4,421,737
Wednesday	3,677,807	2,277,850	4,564,335
Thursday	3,667,126	2,276,408	4,582,144
Friday	3,762,336	2,409,600	4,880,267
Number of stations (1)	400	130	233
Number of tube line	13	4	17
Area (2)	1,572 km^2	718.3 km^2	2267 km^2
Total population (3)	8.63 million	5.3 million	21.15 million
Ridership of Metro	20%	35%	21%
Length of metro lines	402km	182km (MRT+LRT)	465 km

(1) Number of stations is the number of stations with smart-card records generated.
(2) The area of Beijing only counts the area enclosed by the 6th ring road for a fair comparison.
(3) From the World Population Review, http://worldpopulationreview.com/world-cities/ accessed 17 January 2016

Zhong, C., Batty, M., Manley, E., Wan, J., Wang, Z., Che, F., and Schmitt, G. (2016) Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing using Smart-Card Data., **PLOS One**, http://dx.doi.org/10.1371/journal.pone.0149222

From 1 minute intervals to the whole day



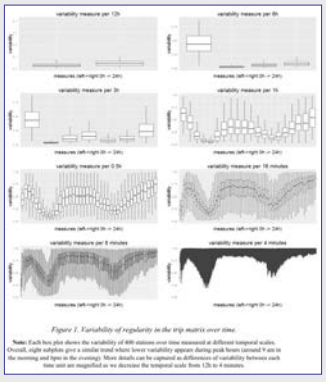
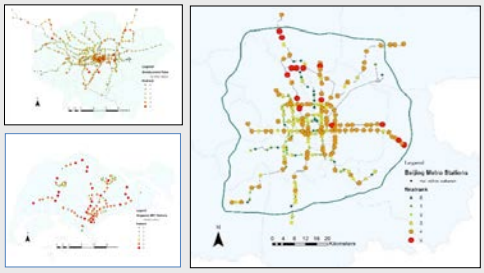
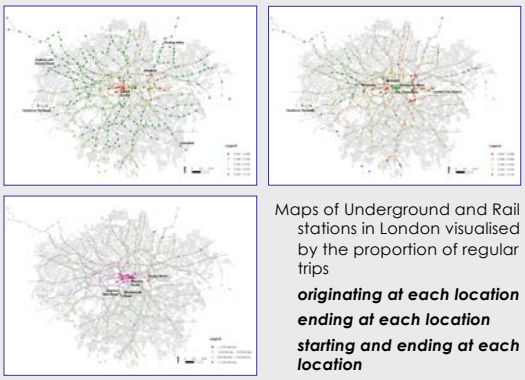


Figure 1. Variability of regularity in the trip matrix over time.

Comparing Variability for different time Intervals over the day

Comparing Variability for different time intervals for Three World Cities: London, Beijing and Singapore





Maps of Underground and Rail stations in London visualised by the proportion of regular trips

originating at each location

ending at each location

starting and ending at each location

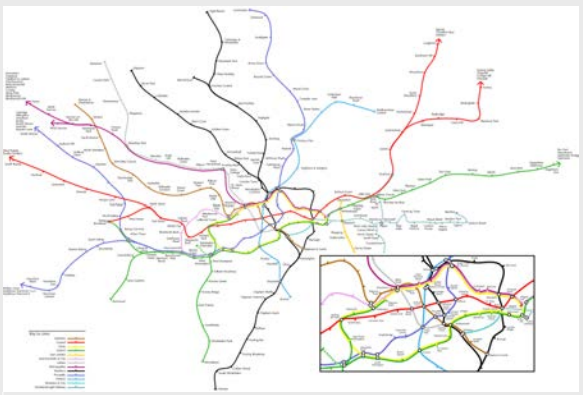
Disruptions – Routine Analysis of Daily Events

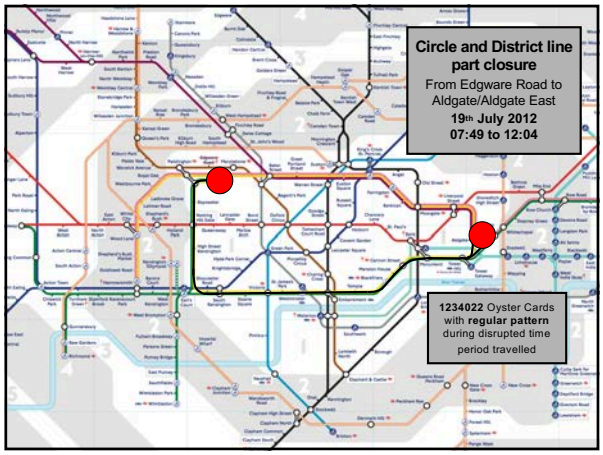
- Behaviours vary across network
- Different areas of network more resilient to disruption, due to available infrastructure and individual ability to change
- But areas of network are naturally closely tied through established usage patterns
- Individual-based analyses provide insight into behaviours underlying macroscopic flows

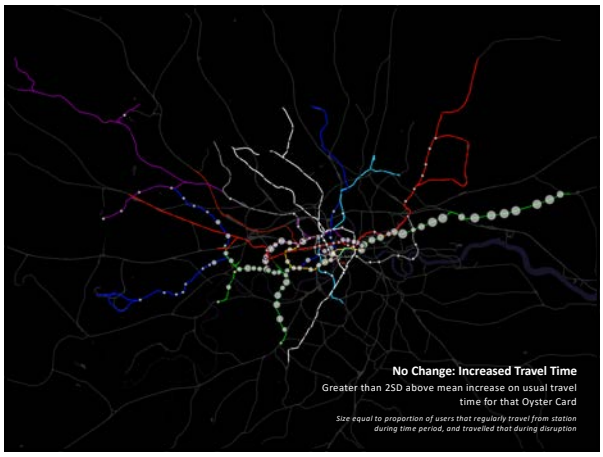
We will look at several kinds of disruption

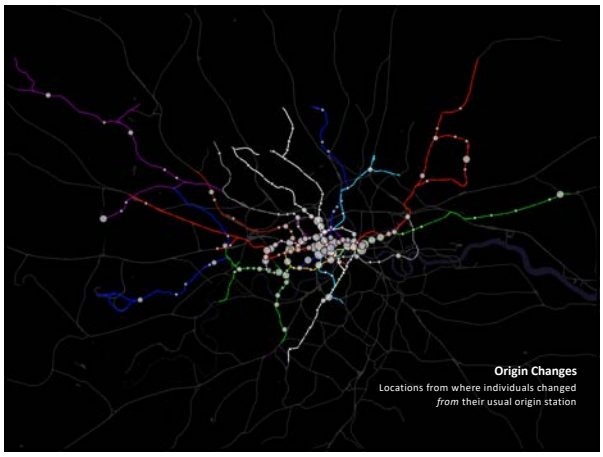
- First hypothetical disruptions simply by examining breaks in the network
- Then an example of the Circle and District Lines which had a 4 hour stoppage on July 19th 2012
- And a Bus Strike in East London and how this shows up in the data
- And typical pattern of delay on all modes visualised for Greater London

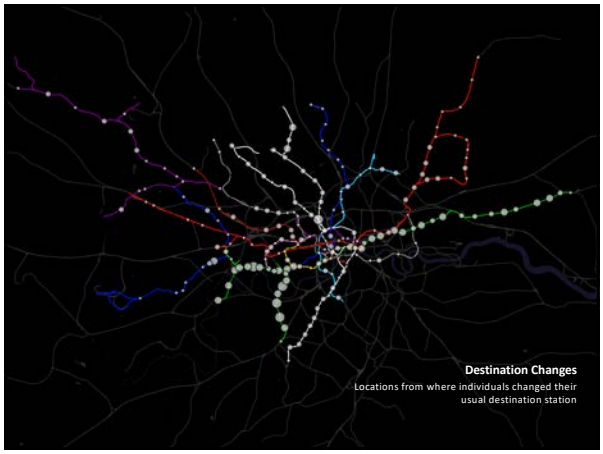






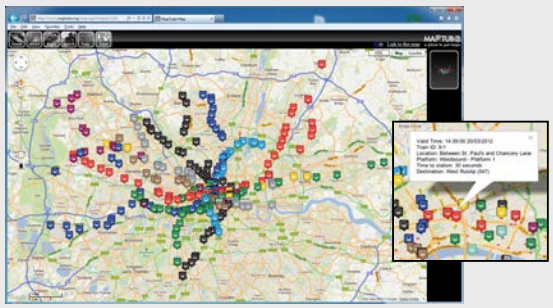




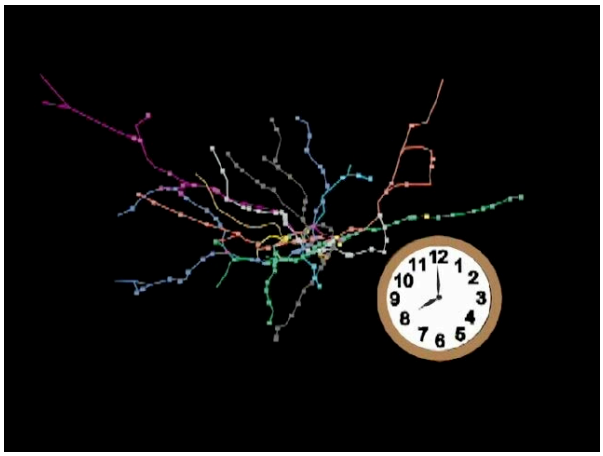


Learning about Mobility from the Data

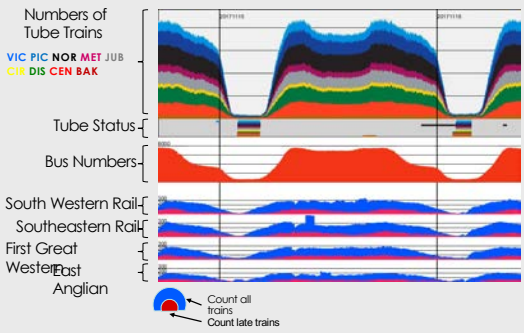
Variable Locational Dynamics of Demand



Hong Kong PolyU: Smart Cities Lectures –June 2019: Big Data, Mobility, and Smart Cities

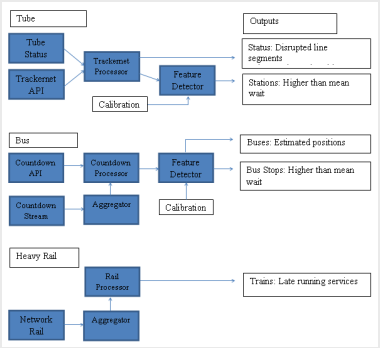


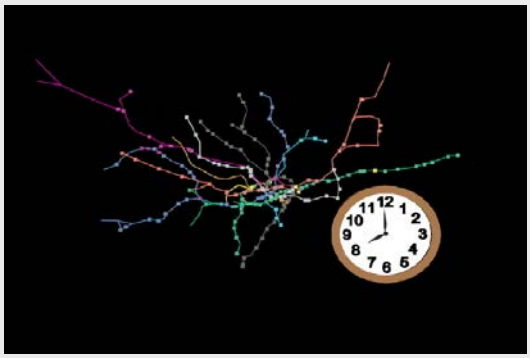
Streams of City Data



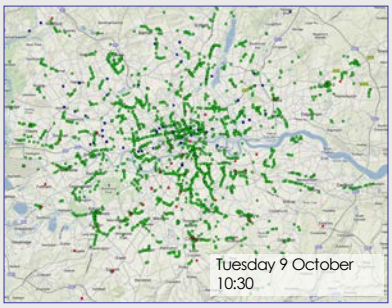
Hong Kong PolyU: Smart Cities Lectures –June 2019: Big Data, Mobility, and Smart Cities

The Public Transport System in Terms of Vehicle Flows





Delays from Tube, National Rail and Bus Fused

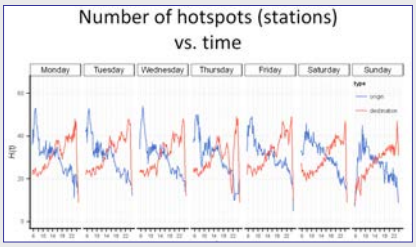


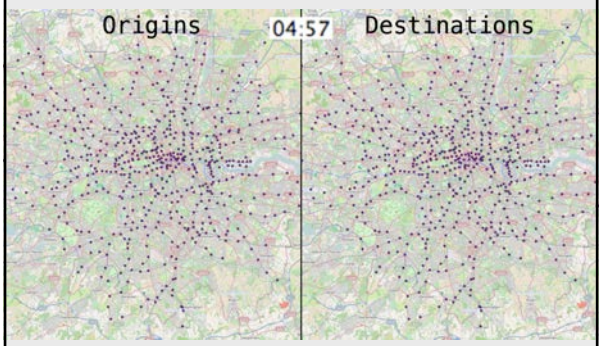
Key

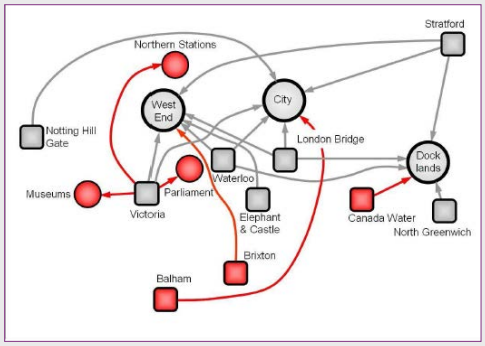
- National Rail more than 5 minutes late
- Tube stations showing a wait time 15% above expected
- Bus stops showing a wait time 20% above expected

Tube delays from the TfL status feed are also plotted as lines

We are also working with the Oyster data on extracting clusters from the travel data using a new method of defining intensity. I will show this as a simple movie of origin and destination intensities as they change over time of day.







Roth C., Kang S. M., Batty, M., and Barthelemy, M. (2011) Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. **PLoS ONE 6(1):** e15923. doi:10.1371/journal.pone.0015923

Lots of Real -Time Data: Bikes, Social Media. Airbnb,

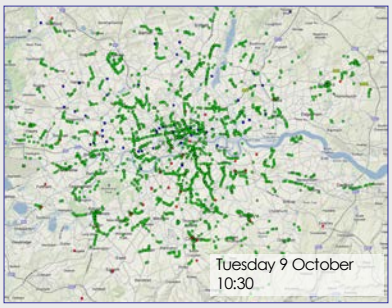
We are doing quite a lot of different exploratory analysis of these new data; and to finish let me just throw out a few of the analyses







Delays from Tube, National Rail and Bus Fused



Key

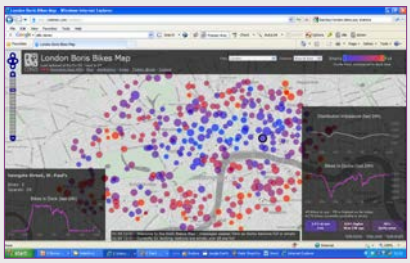
- National Rail more than 5 minutes late
- Tube stations showing a wait time 15% above expected
- Bus stops showing a wait time 20% above expected

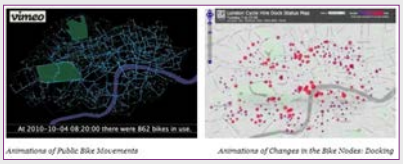
Tube delays from the TfL status feed are also plotted as lines

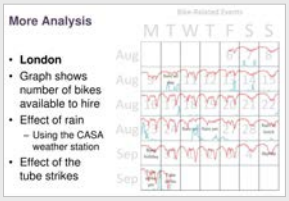
Related Real-Time Data: Bikes, Social Media

A lot of data is now coming online for travel and one of our group Oliver O'Brien has some 97 bike schemes world wide for which he has online data in real time - Bikes Data – 4200 bikes, started Nov 2010, all the data– everything – all trips, all times, all stations/docks











Creating Data: CrowdSourcing

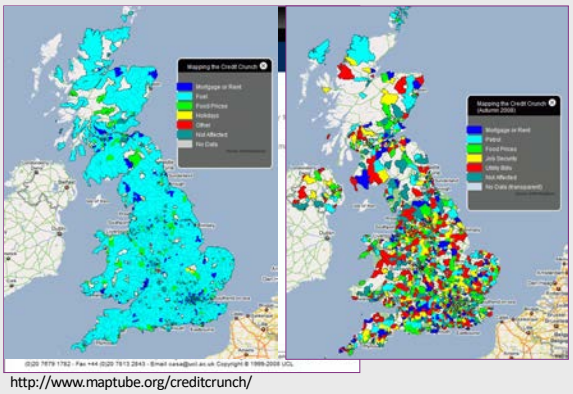
Crowdsourcing is using the power of digital media to create data – this can be done in highly structured ways all the way through to routine ways. They range from online censuses to fully interactive ways of discovering new things online which might be regarded as a form of data

We will look at some online examples. In this we need a broadcast media to make the audience or community aware that we need them to interact and a digital online media to capture their responses – their data



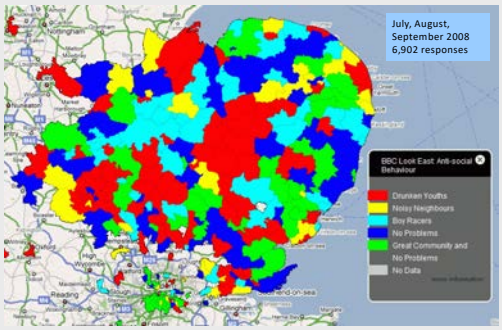
23,475 responses
April, May, June 2008

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



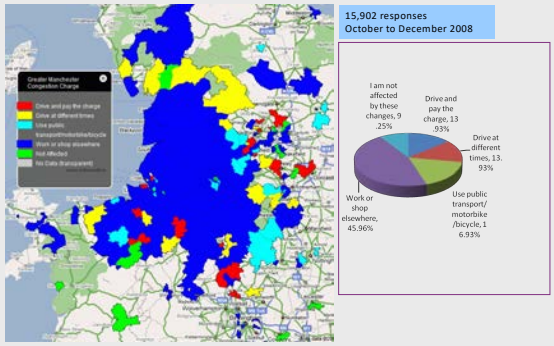
http://www.maptube.org/creditcrunch/

BBC Look East: Anti-Social Behaviour

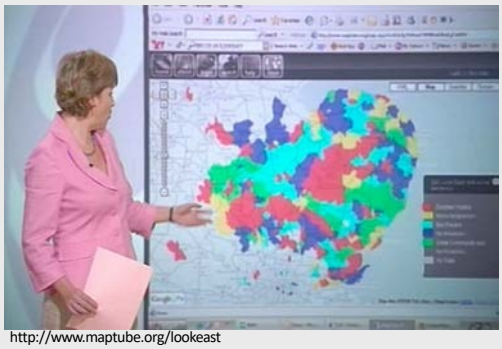


http://www.maptube.org/lookeast

Manchester Congestion Charge

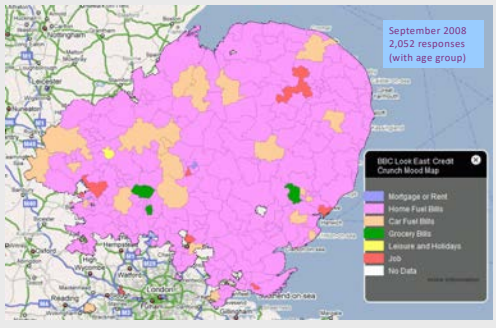


BBC Look East: Anti-Social Behaviour



http://www.maptube.org/lookeast

BBC Look East: Credit Crunch

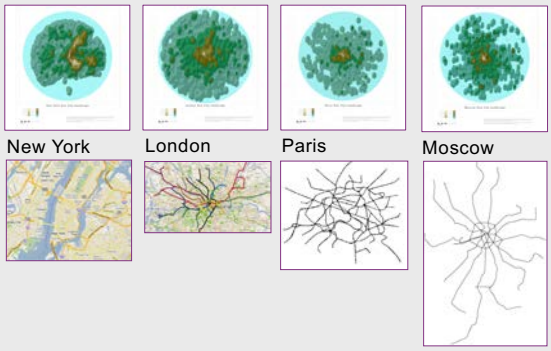


http://www.maptube.org/LookEastCreditCrunch/

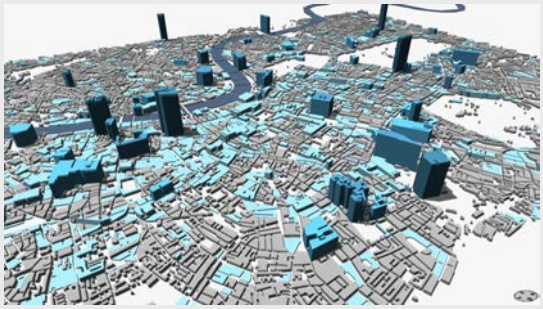
Data, Sensing, Capture, Extraction:

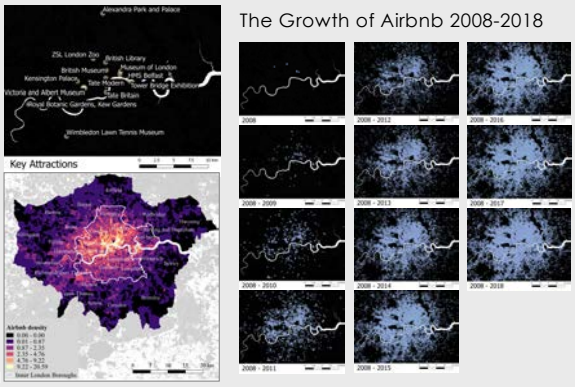


Crowd-Sourcing: Survey Mapper let's you create a survey and mount it on the web; this is part of the BigDataToolkit

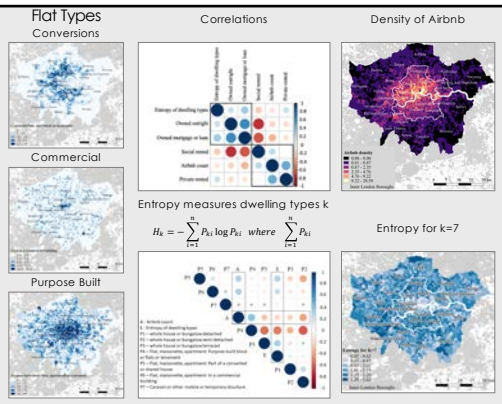








The Growth of Airbnb 2008-2018



Big Data and the City

Editor: Michael Batty
Centre for Advanced Spatial Analysis, University College London

Built Environment

Volume 42, number 3, September 2016

Big data is everywhere, largely generated by automated systems operating in real time that potentially tell us how cities are performing and changing. A product of the smart city, it is providing us with novel data sets that suggest ways in which we might plan better,anddesignmore sustainable environments. The articles in this issue tell us how scientists and planners are using big data to better understand everything from new forms of mobility in transport systems to new uses of social media. Together, they reveal how visualization is fast becoming an integral part of developing a thorough understanding of our cities.



http://www.spatialcomplexity.info/archives/3026

CONTENTS

Editorial: Big Data, Cities and Herodotus
MICHAEL BATTY
Big Data and the City
MICHAEL BATTY
From Origins to Destinations: The Past, Present and Future of Visualizing Flow Maps
MATTHEW CLAUDEL, TILL NAGEL and CARLO RATTI
Towards a Better Understanding of Cities Using Mobility Data
MAXIME LENORMAND and JOSE J. RAMASCO
Finding Pearls in London's Oysters
JONATHAN READES, CHEN ZHONG, ED MANLEY, RICHARD MILTON and MICHAEL BATTY
A Classification of Multidimensional Open Data for Urban Morphology
ALEXANDROS ALEXIOU, ALEX SINGLETON and PAUL A. LONGLEY
User-Generated Big Data and Urban Morphology
A.T. CROOKS, A. CROITORU, A. JENKINS, R. MAHABIR, P. AGOURIS and A. STEFANIDIS
Sensing Spatiotemporal Patterns in Urban Areas: Analytics and Visualizations Using the Integrated Multimedia City Data Platform
PIYUSHIMITA (VONU) THAKURIAH, KATARZYNA SILA-NOWICKA, and JORGE GONZALEZ PAULE
Playful Cities: Crowdsourcing Urban Happiness with Web Games
DANIELE QUERCIA
Big Data for Healthy Cities: Using Location-Aware Technologies, Open Data and 3D Urban Models to Design Healthier Built Environments
HARVEY J. MILLER and KRISTIN TOLLE
Improving the Veracity of Open and Real-Time Urban Data
GAVIN MCARDLE and ROB KITCHIN
Wise Cities: 'Old' Big Data and 'Slow' Real Time
FABIO CARRERA
Collecting and Visualizing Real-Time Urban Data Through City Dashboards
STEVEN GRAY, OLIVER O'BRIEN AND STEPHAN HUGEL

Edited by David Banister and Stephen Marshall with Reviews Editor Lucy Natarajan, *Built Environment* is published four times a year, both print and online. For more information or to order a copy of this issue contact:
Alexandrine Press, 1 The Farthings, Marcham, Oxfordshire OX13 6QD
phone : **01865 391518** fax : **01865 391687**
e-mail : **alexandrine@rudkinassociates.co.uk**
visit : **www.alexandrinepress.co.uk**

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學 | LSGI DEPARTMENT OF LAND SURVEYING AND GEO-INFORMATICS 土地測量及地理資訊學系 | UCL CASA University College London

Thanks

Acknowledgements

Melanie Bosredon, Gareth Simons, Roberto Murcio, Richard Milton, Oliver O'Brien, Stephen Gray, Fabian Neuhaus, Jon Reades, Ed Manley, Chen Zhong, Flora Roumpani & Stephan Hugel

http://www.complexcity.info/
http://www.spatialcomplexity.info/
http://blogs.casa.ucl.ac.uk/

m.batty@ucl.ac.uk
@jmichaelbatty