Institution: City University London

Unit of Assessment: 11 Computer Science and Informatics

Title of case study: Making sense of complex data through innovations in visualisation

1. Summary of the impact

New visualisation approaches have been used to turn complex data into actionable knowledge by:

- *The Ministry of Defence* to establish analytical possibilities for security critical data analysis
- *Transport for London* (TfL) to manage and extend London's successful Cycle Hire Scheme
- *E.ON* to interpret data produced through their modelling and in their Smart Home trial, with a view to understanding electricity consumption and reducing production
- *Leicestershire County Council* (LCC) to develop an evidence base for a sustainable transport plan; record and analyse the locations of locally valued green spaces; capture local knowledge about flooding events; monitor performance of children's centres; present the results of a survey on service quality and accessibility to citizens; undertake a £100M budget consultation and embed data in decision-making processes to inform policy
- *Willis* to understand and assess windstorm risk, communicate the complexity of risk to clients and manage risk across their global offices through a new software system.

These applications of new visualisation methods have had impact on the environment, economy, defence and security, society and public debate. In each case users of our methods report on their positive impact as we help them identify visualisation possibilities, understand their data and use this knowledge to inform their activity. In many cases our work has resulted in important insights, improved exploitation of data and further investment in visualisation with organisational implications in terms of using data for intelligence.

2. Underpinning research

Professionals, policy-makers and the public struggle to interpret the large amounts of data to which they are exposed. We address this problem and take advantage of the data deluge by developing new interactive visualisation software, techniques and algorithms. These are designed and evaluated in the context of client need in our problem-driven research. They include:

- A suite of **methods** that combine principles of spatial cartographic layout with non-spatial information in ways not previously considered in information visualisation. Examples include spatial treemaps (Wood and Dykes, 2008); origin-destination maps (OD maps - Wood *et al*., 2010) and flow maps (Wood *et al*., 2011). In addition, our theoretical framework for exploring the visualisation 'design space' (Slingsby *et al*., 2009) has allowed us to create graphical depictions of complex data tailored to particular scenarios and specific lines of enquiry.
- A **design methodology** for establishing analytic requirements that we then address by selecting from visual design possibilities (e.g., Lloyd and Dykes, 2011; Dykes *et al*., 2010; Walker *et al*., 2013). It uses rapid data-rich digital prototyping as part of an iterative requirements-design-feedback process. This makes it particularly suitable for contexts where complex data are used and requirements are somewhat open – as is the case with new technologies and new data. Unlike prior elicitation methods our approach is specifically tailored towards contexts in which rich and complex data are key, with 'data sketches', which are software prototypes developed to show real data rapidly and in context through 'chauffeuring' as part of the design process.

Our research is driven by our engagement with data owners in a range of application domains. For example, our engagement with a local authority that needed to consider population, geography and administrative hierarchy concurrently led to our development of spatial treemaps.

The research reported here was undertaken at City University London between 2006 and 2013 by Professor Jo Wood, Professor Jason Dykes, Dr Aidan Slingsby and Dr David Lloyd in collaboration with industry and government.

3. References to the research

The research is published in the leading peer-reviewed journals. The team has received 5 awards for publications relating to the impact described, with a further 9 received for other outputs. The research has been supported by over £770,000 of funding from national funding councils (EPSRC, ESRC, JISC - £344k), industry (Willis, E.ON - £384k) and government (DCLG, LCC - £45K).

Wood J. & Dykes J. (2008) Spatially ordered treemaps, *IEEE Trans. V&CG*, 14(6) 1348-1355.

10.1109/TVCG.2008.165 [Based on work awarded GIS Research UK 2008 / Association of Geographic Information Best Paper]

Slingsby A., Dykes J. & Wood J. (2009). Configuring hierarchical layouts to address research questions. *IEEE Trans. V&CG* 15 (6), 977-984 10.1109/TVCG.2009.128 [IEEE Infovis 2009 Honourable Mention]

Wood J., Dykes J., Slingsby A. (2010). Visualization of origins, destinations and flows with OD maps, *Cartographic Journal* 47(2) 117-129 10.1179/000870410X12658023467367 [Based on work awarded GIS Research UK 2009 / Association of Geographic Information Best Paper]

Lloyd D. & Dykes J. (2011). Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE Trans V&CG*, 17(12), 2498–2507 10.1109/TVCG.2011.209

Wood J., Slingsby A. & Dykes, J. (2011). Visualizing the dynamics of London's bicycle hire scheme. *Cartographica*, 46(4), 239-251 http://openaccess.city.ac.uk/538/

Walker R., Slingsby A., Dykes J., Xu K., Wood J., Nguyen P., Stephens D., Wong W. & Zheng, Y. (2013). An Extensible Framework for Provenance in Human Terrain Visual Analytics. *IEEE Trans. V&CG*, 19(12) 2139-2148 10.1109/TVCG.2013.132

4. Details of the impact

1. Defence Science and Technology Laboratory (DSTL) uses science and technology for the defence and security of the UK. We have helped DSTL develop new possibilities for analysing the kinds of complex, conflicting and unstructured data collected in areas of conflict such as Afghanistan and Iraq.

Our architecture, software prototypes and other research outputs have enabled DSTL to establish possibilities, identify analytical requirements and develop internal (confidential) software capability. Leo Borrett, Capability Lead at the Counter Terrorism Science and Technology Centre confirms that: "*Through working with City … MOD's understanding of its requirements for multi-faceted data visualisation is sufficiently mature that we are able to start placing external work for software development to support our intelligence analysis applications. The concepts presented … are now central to our thinking when planning for future visual interfaces for security critical data analysis.*"

Our techniques and prototype visualisation applications are deemed "*exploitable*" within DSTL's MAMBA framework for analysis. (MAMBA is a high-level RAD (rapid application development) Web Applications framework.)

2. Transport for London (TfL) is responsible for London's transport system. Our visualisation work with the Cycle Delivery Planning team has enabled them to iteratively develop knowledge of the complexities of cycling behaviour including when and where cyclists start and end journeys and how different types of cyclist behave. This helps TfL provide bicycles in the right places at the right times for users of the scheme with broad benefits: "*fitness, enjoyment and easy travel for millions, cleaner air and less traffic for all*" (*http://bit.ly/19GacTQ*).

Our visualisation of the journeys made since the start of the scheme (over 20 million) detected important, previously unknown patterns, including responses to station closures and differences in behaviour between locals and commuters, cyclists in west and east London and males and females. For example, female cyclists select quieter parts of the city for their journeys and are less likely to cross the Thames than male cyclists.

Our work has been used in the day-to-day running of the scheme and the information it has uncovered about general cycling patterns supports wider ambitions and longer term planning. It is contributing positively to "*the main cross-London physical legacy of the 2012 Olympic Games*" on which £913M will be spent over the next decade (*http://bit.ly/19GacTQ*).

Peter Wright, Senior Cycling Delivery Planning Manager for TfL, confirms that this work is being used to (1) inform decision-making: "*Your visualisations have been so useful for getting the message across to others in TfL and borough planning officers*"; "*the findings you have made so far are already influencing some of our decision-making*"; (2) inform policy: "*your customer classification and analysis of geographic trends has informed phase 3 of the scheme's expansion into south-west London and intensification of the existing area*"; and (3) secure the Mayor's vision: "*The analysis of gendered motivations and barriers to using the scheme tell us something about broad cycling behaviour in London. This substantial evidence base will help us secure sustained political and financial support to the ambitious plans set out in the Mayor's 2013 Vision for Cycling*

in London".

3. E.ON is a leading energy provider in the UK, supplying 5 million people with energy. The New Technology team has significantly improved its understanding of home energy consumption by working with us to visualise data collected through a trial of Smart Home technology installed in 100 houses in Milton Keynes. They have also established new and effective ways of visually analysing these kinds of large dynamic data sets by participating with us in creative visualisation development work. This has enabled the team to see previously unknown patterns in the Smart Home data and created the potential for using data more effectively in their organisation.

In evaluating our visualisation solutions, E.ON data analysts made the following comments: "*very powerful and very useful*"; "*18 million data points! [It] is just impossible for us to get our head around the real value that is contained in that*"; "*it gives us a whole new way of analysing people*"; "*you could spend months searching the data for insights but this just points you straight at it*"; "*the [Smart Home] project will be better for it*".

Our work uncovered previously unknown characteristics in the data, which persuaded E.ON to use the Smart Home data in their decision-making. This is evidenced again in the evaluation: "*I wouldn't be able to spot the problem before I saw this graph*"; "*I didn't expect to see these patterns*"; "*before I thought the trial data could not be used due to errors and outliers. The visualisation showed me that you can use this data and detect different patterns and user behaviour*". These reactions reflect how our research can provoke a change in approach to data collection and use. The visualisations led to the discussion of new approaches to managing peak demand and the establishment of a new long-term plan with the objective of flattening peak consumption and reducing energy production. Our analytical work led to direct ideas about offsetting and delaying consumption associated with particular household devices such as freezers, washers and dryers.

4. We have worked with **Leicestershire County Council** (LCC**)** to develop techniques and design methods to support decision-making, improve services and planning and inform citizens.

Andy Robinson, Assistant Chief Executive at LCC reflects: "*We're making better decisions, more informed policy and better use of our data across the organization ... The City visualisation work has been fundamental in helping us do so.*"

Our techniques were developed with and used by LCC. Spatial treemaps and Origins and Destinations (OD) maps were used to interpret commuting patterns and identify areas for LCC intervention in the Local Transport Plan (LTP) evidence base. The LTP is a major report required by the Transport Act 2000 to encourage effective delivery of high quality transport planning. It aims to contribute to economic growth, emission reduction, quality of life, safety and health. It also promotes equality and opportunity and defines the basis for monitoring transport performance. Our visualisation contributed directly to LCC's sustainable transport policy (*http://bit.ly/9r98dQ*). Steve Rothwell of LCC's transport policy team confirms that the maps in the evidence base helped the team understand the data in developing the plan: "*the graphics have been really useful and have helped us to highlight possible areas of intervention*" and that the OD maps enabled analysts to "*genuinely get a sense of commuting patterns in Leicestershire*".

We have also developed and used novel visualisation systems with and for LCC:

Place Survey (*http://bit.ly/gRVKkb*) delivers information to citizens on attitudes to quality of life and services across the county in more detail than ever before, through a completely new approach using personalisable, exploratory graphics. It also helped to inform an LCC initiative to improve access to services through closer work between partnerships: "*[PlaceSurvey] enabled the place survey to be easily interpreted so that we could understand issues for specific demographics within the locality at ward level – which fed into the proposal for an integrated model of service delivery*".

Green Spaces enabled LCC to run a policy-forming public consultation to capture (*http://bit.ly/18nsPty*) and analyse (*http://bit.ly/hhdODT*) the parks and open spaces valued by citizens in response to a government initiative to safeguard the environment. Our web-based interface helped raise awareness of the issues - with local campaign groups encouraging participation - and made the evidence base available to planners and the public. The work was commended as "*really innovative*" by the council leader in a December 2010 speech to Council where he urged district councils to use it to "*actually look at what people in Leicestershire are saying*". It has generated local discourse that has enriched the public consultation (*http://bit.ly/1bFckNY*). Harborough district, Melton district and the Fosse Villages Neighbourhood Plan group use the software for identifying green spaces for protection through local and neighbourhood plans.

Flood Spaces (*http://bit.ly/182NEuy*) is helping LCC build a comprehensive picture of flood risk through public engagement. Doing so meets a statutory requirement arising from the Flood and Water Management Act 2010 to assess the flood risk within the County. The crowdsourcing that our work facilitated captures local knowledge in advance of flood incidents. As many flood incidents go unreported and a recent flood in Market Harborough caused damage in excess of £1M, this data is valuable in terms of economic savings and environmental protection.

Our relationship with LCC has resulted in a wide and pioneering uptake of visualisation that they describe as supporting "*situation-driven data use*". Visualisation approaches introduced by and developed with City are being used in numerous departments (Performance; Research; Public Health; Museums; Climate Action Team) for various purposes (scrutinising expenses and mileage to make savings and reduce carbon footprint; considering social fund expenditure by area; analysing museum income). Performance monitoring of children's centres is an important example whereby services have improved through more effective use of data: "*We failed [OFSTED] inspections in the past - but we don't now due to the data visualisation work*" (Robert Radburn, Research and Insight). The relationship between LCC and City is reciprocal: the needs of LCC have inspired many of our transferable designs: e.g., spatial treemaps and OD maps.

As a result of this knowledge transfer, LCC increasingly adopts visual approaches to data analysis that use techniques and methods established through our collaborative work. For example, the current budget consultation at LCC will radically change the council as £110M of savings must be delivered to achieve a 30% cut to the budget. Visualisation is being used to inform this consultation: the three key decision-makers receive visual summaries of consultation responses, updated each hour, on their desktop. These details informed the initial savings required for October 2013 and will be used in subsequent spending reviews.

The work has been funded through short-term consultancy, Research Council fellowships (in both directions) and research grants (e.g., Department for Communities and Local Government *Timely Information for Citizens*). It has resulted in the establishment of information visualisation at the core of the organisation's use of data, which is increasingly being used to inform decisions (*http://bit.ly/13Px67X*). This approach is "*opening up data to improve policy debates [by] enabling policy officers to analyse and discuss data on demand*" and resulting in "*better conversations internally and better service from Research and Insight*" in an organization that employs 15,000 staff to support 600,000 residents through a range of services in the context of significant pressure on its £300M budget.

5. Willis, one of the world's largest insurance brokers, have improved their understanding and communication of risk in the insurance industry through the use of our visualisations. For example, we furthered understanding of windstorm risk by presenting complex and uncertain windstorm model outputs in an accessible and interpretable way. This work was long-listed for Lloyd's 2011 Science of Risk Prize and won the prestigious 2011 IEEE Discovery Exhibition award for industry impact (*http://bit.ly/10WuYKs*).

Willis use our visualisation approaches to proactively manage risk and communicate its complexities to their clients at their facilities across the globe. For example, Atlas - a system supporting this activity - incorporates work undertaken through a City MSc project.

Willis funded Slingsby's post-doctoral position for two three-year terms through the Willis Research Network (WRN). We are core members of what is the world's largest collaboration between public science and the financial sector, developed to "*improve resilience by integrating first class science into operational and financial decision-making*" (http://bit.ly/191DGHf).

5. Sources to corroborate the impact

Robert Radburn - Research and Insight, **Leicestershire County Council**, County Hall, Glenfield, Leicester LE3 8RA. *Robert.Radburn@leics.gov.uk*

Chris Utting - New Technology Consultant, Energy Infrastructure & End Use, **E.ON**. 02476 181504; 077930 99436. *Chris.Utting@eon.com*

Peter Wright - Senior Cycling Delivery Planning Manager, Surface Transport, **TfL**, Palestra 11Y4, 197 Blackfriars Road, London SE1 8NJ. 020 3054 7134; *PeterWright2@tfl.gov.uk*

Nigel Davis - Managing Director, Platforms and Delivery, Global Analytics, **Willis Group**, 19th Floor, The Willis Building, 51 Lime Street, London, EC3M 7DQ. 020 3124 8959; *davisn@willis.com*

Leo Borrett - Capability Leader, Counter Terrorism Science and Technology Centre, Room 202-E, Building 5, **DSTL** Porton Down, Salisbury, Wiltshire, SP4 0JQ. 01980 658865; lcborrett@dstl.gov.uk