The Urban Big Data Centre: a Licensing Case Study

Choose the Right Rights, Use the Data Right

2 November 2017
What we do
We are UK’s national data service for urban research

UBDC Portfolio

Data Products
Data Service
Research

Interdisciplinary perspectives –
Urban Social Science, Data Science and Engineering

- Urban studies
- Urban planning & policy
- Statistics
- Economics
- Computer science
- Education
- Geography
- Civil engineering

Outreach/KE
Training

25 Co-Principal Investigators
300+ stakeholder organizations
Major Achievements

• Developed strong relationships with data owners, acquiring and producing a range of popular datasets
• Built an IT and governance infrastructure that accommodates a range of data, users and usage
• Established a growing base of users from several domains, including academia, commercial and government sectors
• Engaged with stakeholders locally and internationally
• Delivered an acclaimed programme of training and capacity building
Key Data Sources

**IN-HOUSE DATA PROGRAMMES**
- IMCD
- SUDS

**LOCAL AUTHORITIES & GOVERNMENT**
- Transport Scotland
- Registers of Scotland
- UCAS
- Glasgow City Council
- Renfrewshire Council

**COMMERCIAL ORGANISATIONS**
- STRAVA
- Zoopla
- Twitter
- Nestoria
- BlueSky

**Institution Logos**
- ESRRC Economic Social Research Council
- University of Glasgow
- Urban Big Data Centre
- An ESRC Data Investment
Core Challenge

• **Acquiring and Licensing data** –
  • How to incentivise organisations to share their data, and to permit further sharing to wider audiences?
  • Incompatibilities with business models, concerns over reputational harm (e.g. bad data, or exposure to criticisms based on analysis of data) or just lack of resource or internal influence slow the process
  • Even *physically* acquiring data can be non-trivial
  • How to manage multiple bespoke licensing agreements in a single data service platform? How to develop our own commercial offer?
  • How best to enforce usage that is conformant with restrictions, but still provide a service that supports good research? Not all users are the same.
Managing sensitivities

• Personal data
  • Data Protection Act / General Data Protection Regulation

• Commercial data
  • Do data owners have existing income streams associated with providing access to data?

• Data that may be used in research that casts the owner in a poor light
Long term availability

• UBDC currently funded to January 2019
• Our default option at that time is to deposit in UKDS
• How to ensure compatibility of our licensing approach (as well as technical approaches?)
• Even bipartite relationships are complex to negotiate
Open and Safeguarded data

• Open data
  • Accessible via public (CKAN) portal, no restrictions on sharing or reuse

• Safeguarded
  • Subject to licensing restrictions
    • UBDC EULA / Standard Data Sharing Agreement
    • Bespoke Sub-license agreement
    • Access restricted based on negotiations with data owners
      • Physical limitations on access
      • Non-commercial / research only?
Licensing

• Prospective users of UBDC’s data services are expected to sign our overarching Access Agreement, as well as accept the terms and conditions of any given dataset.

• Terms vary, e.g., some data sets can be used for commercial research, and others are more restrictive

• Liabilities?

• Requirements re: demonstrating conformity, new data products etc
Controlled data service

• We work with the *electronic Data Research and Innovation Service (eDRIS)*, to provide researcher access to controlled data.

• Controllers of **personal data** (which is data that may identify an individual living person, or could do so when combined with other information) may agree to permit others to have access for specific research purposes where there is a public interest case for them to do so.

• In most cases we do not acquire such data, but can help **approved researchers** negotiate access to a **secure environment** to work with controlled data.
Data Showcase - Integrated Multimedia City Data

• The UBDC's Integrated Multimedia City Data project has gathered a unique and exciting collection of data looking at urban life in Glasgow.

• A range of data collection methods including:
  • A household survey and travel diary
  • Sensor data collected from individual subjects who wore life-logging and GPS devices
  • Information retrieval exercises looking at textual and multimedia data sources online
  • Satellite data

• An integrated and linked picture of urban life in Glasgow, attained from multiple perspectives
Core Challenge

• De-identification of data / resolving privacy concerns
  • Demands understanding of privacy legislation, ethical considerations, and of course the characteristics of data and the methods used to collect it, often requiring too the input of colleagues within information compliance services.
  • Supporting resources are limited, risk exposure is significant and aligning user expectations and data requirements with appropriate data and release paradigms is often challenging.
  • GDPR magnifies the risk and re-raises the bar