

Developing Smart Statistics for Urban Mobility: Challenges and Opportunities

Konstantinos Ampountolas (Senior Lecturer; Co-I UBDC)

Andrew McHugh (Senior Data Science Manager UBDC)

Vonu Thakuria (Founding Director UBDC)

Urban Big Data Centre, University of Glasgow, UK



The Urban Big Data Centre

“Promoting innovative research methods and the use of big data to improve social, economic and environmental well-being in cities”

- **UK Government** (Economic and Social Research Council) funded
- Operating a **research-led data service**
 - Data infrastructure and collections
 - Priority research strands: transport & mobility; neighbourhood, housing & environment; education, skills & productivity; big data & urban governance
 - Combining social science research with data analytics and computing science
- **Overall aims:**
 - Achieve public policy impact
 - Critically evaluate role and value of big data and urban analytics
 - Enhance data and methods

The Urban Big Data Centre (<http://ubdc.ac.uk>)

Data service/catalog: <http://ubdc.gla.ac.uk>

Search Your Data

eg. Gold Prices



Popular Tags

population

census

demography

Welcome to the UBDC Data Portal

Here you can access a wide range of urban-related data, covering topics such as commercial, governmental, transportation, social media data and more. We will continue to expand the amount of data available, so be sure to check back for future updates.



Urban
Big
Data
Centre

E · S · R · C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

UBDC Data Portal statistics

1k

datasets

4

organizationsgroups

0

0

related items



University
of Glasgow

Tags

population (109)

census (57)

demography (48)

employment (45)

projections (44)

housing (43)

SUDS (41)

transport (39)

lambeth (37)

health (37)

University
Glasgow

An ESRC Data
Investment

Context

Technological	<ul style="list-style-type: none"> Information generation and capture Data management Data processing Archiving, curation and storage Dissemination and discovery Algorithms / AI / Machine Learning 	Theoretical and epistemological	<ul style="list-style-type: none"> Understanding metrics, definitions, and changing ideologies and methods to solving domain problems Determining validity of approaches and limits to knowledge from data-driven approach Information paradoxes (Jevons paradox), user equilibrium vs system equilibrium
Methodological	<p>Data preparation</p> <ul style="list-style-type: none"> Information retrieval and extraction Data linkage/information integration Data cleaning, anonymization & quality <p>Data analysis</p> <ul style="list-style-type: none"> Methods to analyse domain challenges Uncertainty, biases and error propagation 	Political economy	<ul style="list-style-type: none"> Data entrepreneurship, innovation networks and power structures Value propositions and economic implications Data acquisitions strategies, access and governance framework Privacy, security and trust management Responsible innovation and emergent ethics

Urban data context – data sources

Urban Big Data	Examples
Sensor systems	Environmental, water, transportation, building management sensor systems; connected systems; Internet of Things
User-Generated Content	Participatory sensing systems, citizen science projects, social media, web use, GPS, online social networks and other socially-generated data
Administrative (governmental) Data	Open administrative data on transactions, taxes and revenue, payments and registrations; confidential person-level microdata
Private Sector Data	Customer transactions data from store cards and business records; fleet management systems; usage data from utilities and financial institutions; product purchases and terms of service agreements
Arts and Humanities Data	Repositories of text, images, sound recordings, linguistic data, film, art and material culture, and digital objects, and other media
Hybrid Data Sources and Synthetic Data	Linked data including survey-sensor, census-administrative records

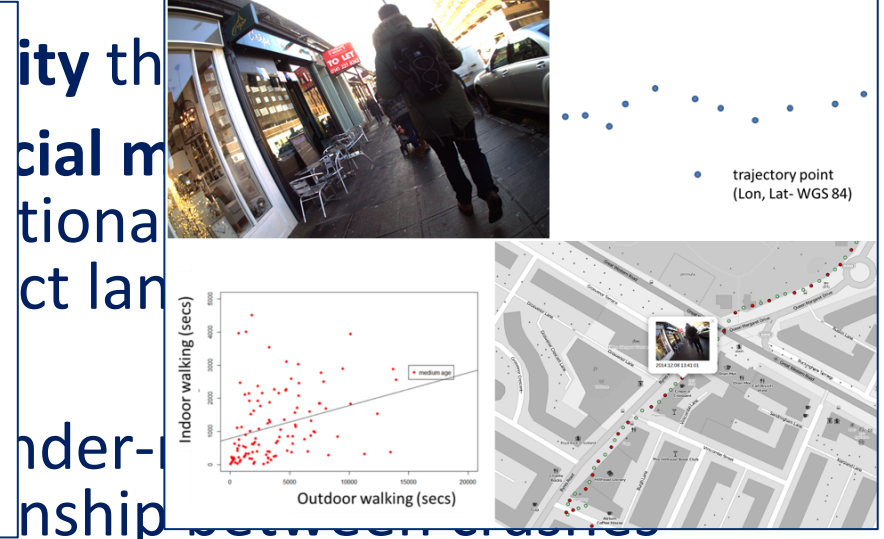
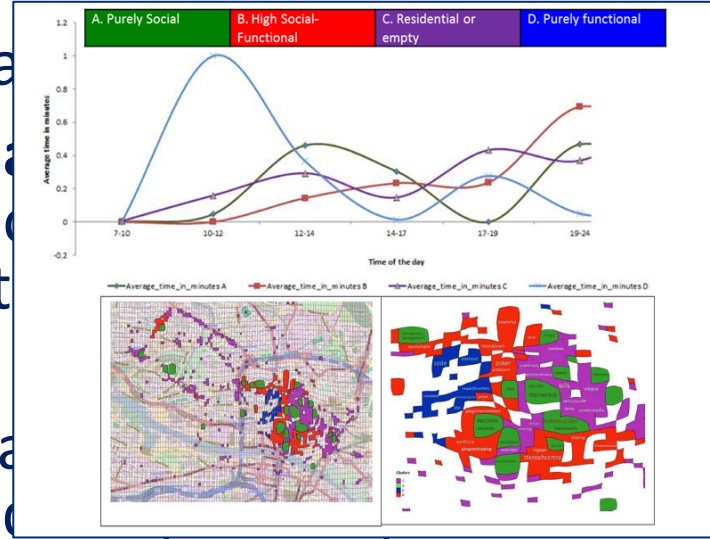
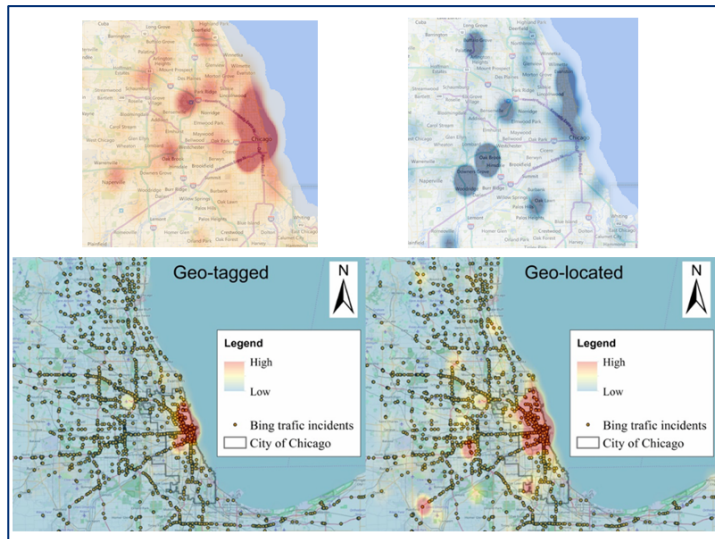
Thakuria, P., N. Tilahun and M. Zellner (2017). Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery. In *Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics*, Springer, NY, pp. 11-48.

Urban data context – research & methods

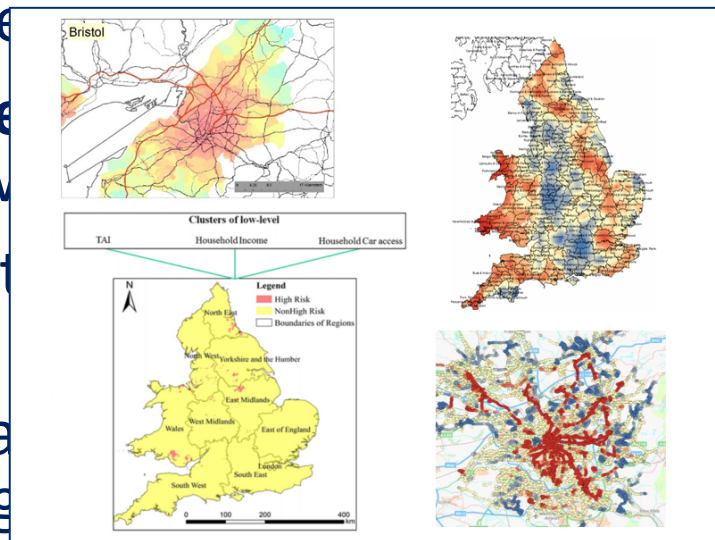
Rich strands of urban analytics within **urban mobility** theme

- Urban metabolism – **real time analytics** using **social media** and **GPS data** to identify spatio-temporal activity clusters (functional usage / stay duration...) and semantically annotated to connect land use PoI and transport networks
- **Geolocalisation** of social media data – identify under-reported phenomena such as road traffic incidents, and explore relationship between crashes and crime
- **Wearable sensors** combined to show mobility patterns and behaviours (indoor walking; social exclusion; travel modality)
- Transport poverty – relationships with labour markets and changing nature of work - **small area transit availability indicators**
- Active travel – using datasets such as **Strava Metro**, validating and informing infrastructural investments

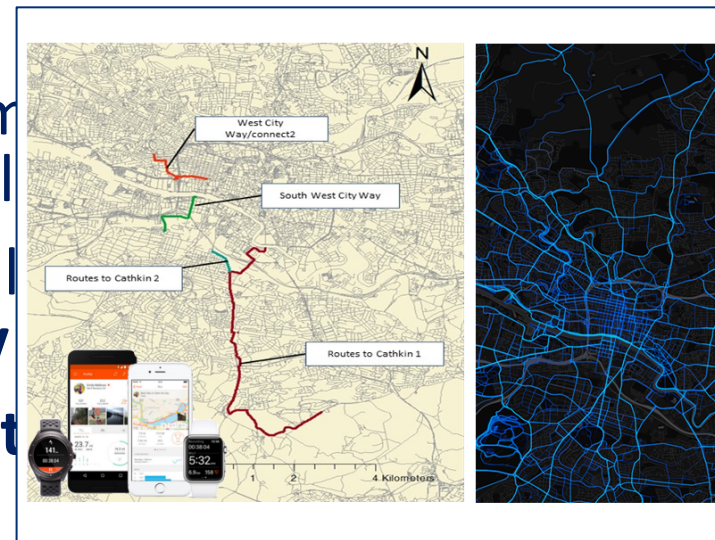
Urban data context – research & methods



- and crime
- **Wearable** (indoor walking)
- Transport of work -
- Active transport informing



low mobility
travel
with
ability
as
nts



behaviours
nature
and

Key challenges



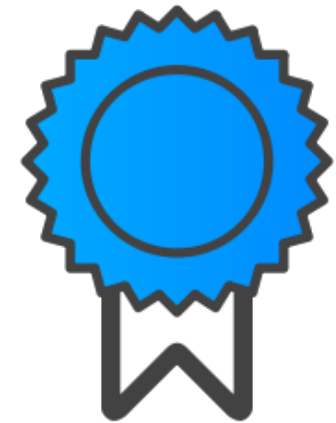
Skills



Technology



Data sharing



Trust

Skills, knowledge and team composition

Multiple academic fields

- Urban studies
- Transport and spatial planning
- Engineering
- Computing Science

Subject specialisms

- Spatial and statistical analytics
- Computing Science:
 - Information / data management
 - Information retrieval
 - Human computer interaction
- Economics, law and information science



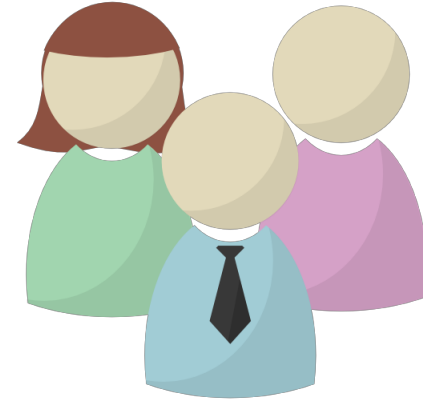
Skills, knowledge and team composition

Skills

- Science of sensors, including remote sensing systems
- AI / machine learning
- DB management / administration
- Data visualisation
- GIS spatial analysis
- Information governance

Team composition

- Domain professionals
- Information scientists
- Statistical analysts
- Legal and ethics
- Consumer needs specialists
- Communications & outreach
- Business modellers

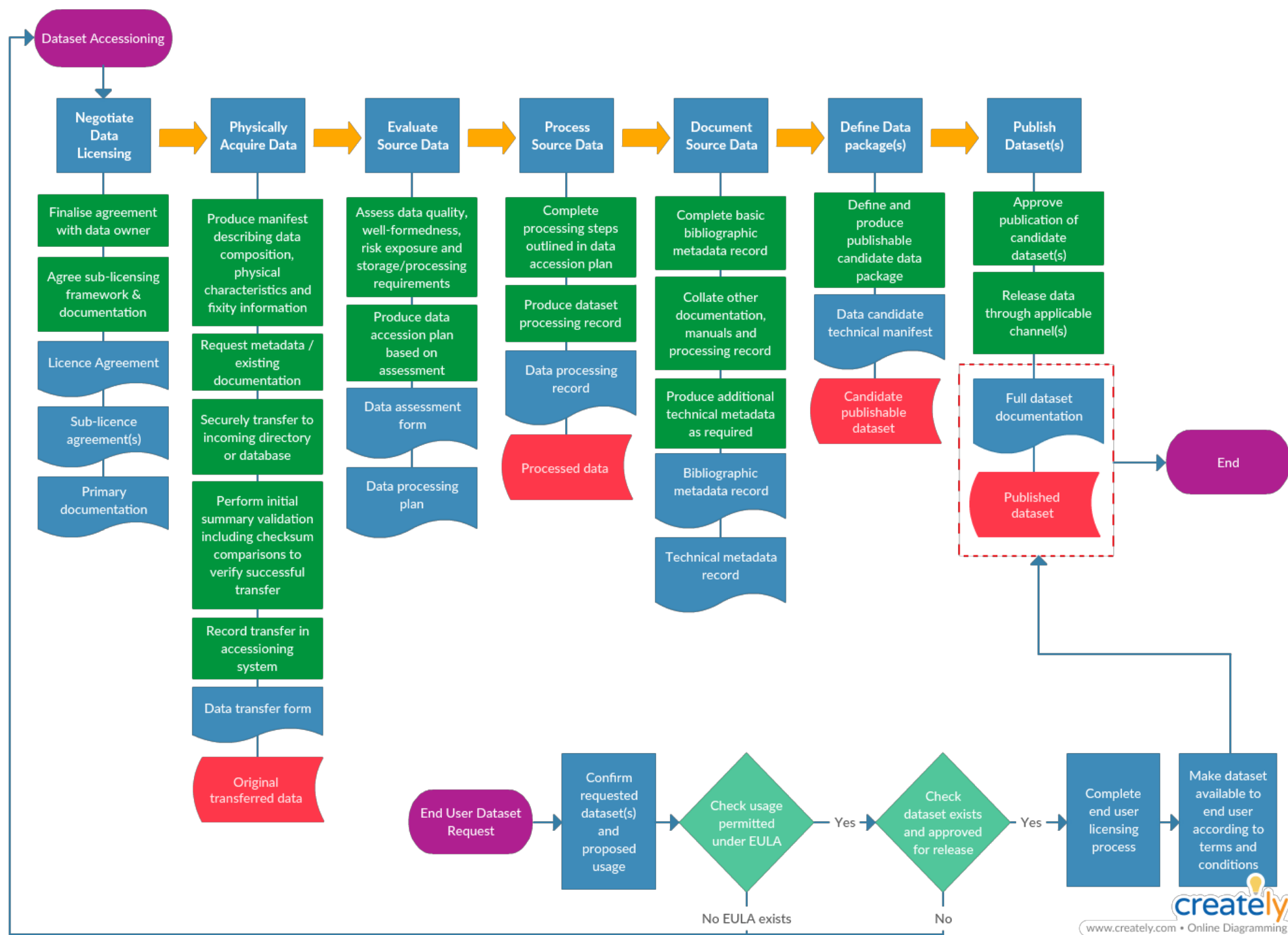


Successful teams learn from each other, listen to needs, are open to new ideas, and are constantly seeking to collaborate.

Technology and structure

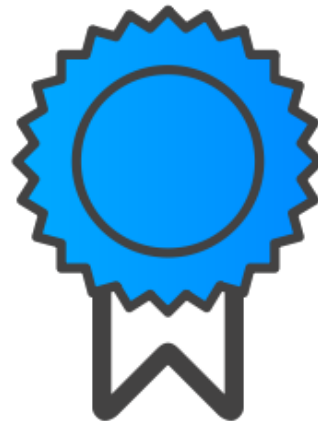
- Significant levels of capacity-building
- Partnerships with academics, industry and local governments
- Fit-for-purpose technological and methodological approaches
- Data standards for harmonisation across countries
- Methodological/algorithmic standardisation
- Having national champions and also local champions to highlight the importance of smart statistics and demonstrate value to key decision-makers and the public – peer review role?
- Peer-to-peer networks to establish collaboration and community-based learning
- Having approaches to query and mine the data in an exploratory sense to understand emerging trends
- Meaningful and impactful derived data and analysis, and proactively demonstrating public good
- Public engagement and informing public of benefits and risks of data (especially necessary when others are now providing critical data)





Trust and validation

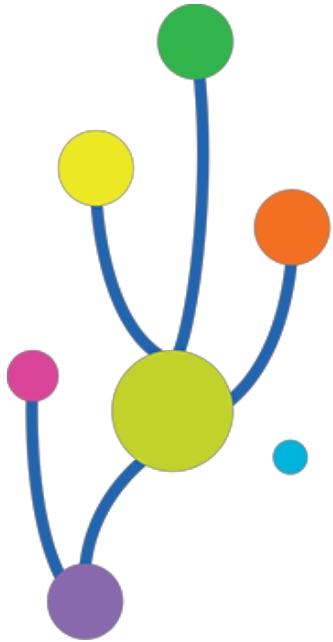
- Expert evaluation/recognised accredited authority/review and oversight of **trust across different sectors**
- Seek **ground truth data** – easier said than done but important to have as aspirational goal
- Soft systems approaches from a methodological perspective to derive weights on results from different methods or different analysts
- Novel methodological approaches to assess and **capture uncertainty** at each stage of data to output lifecycle
- Conduct extensive **sensitivity analysis** and **simulations** to understand behaviour of **statistics and indicators** from “black box” algorithms at every stage of the **parameter tuning** or critical junctures – validations / evaluations should occur at these stages not just with the final product
- **Closing the loop!** Closed-loop feedback statistics





University
of Glasgow

Thank you!



Urban
Big
Data
Centre

An ESRC Data
Investment

**INSPIRING
PEOPLE**