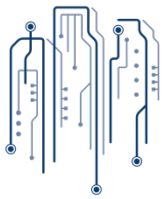
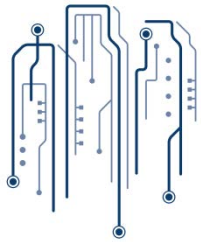
The data privacy, data protection and data security implications of smart cities and urban big data

Prof. Rob Kitchin
Maynooth University

Geo Ethics, Twente
14th March 2016



The Programmable City



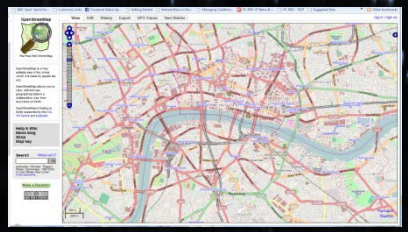
Data and the city

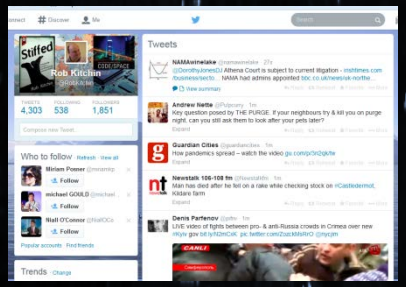
- Rich history of data being generated about cities
- Long had *data-informed urbanism*
- Being complemented and replaced by *data-driven, networked urbanism*

Urban big data

- Directed
  - Surveillance: CCTV, drones/satellite
  - Scaled public admin records
- Automated
  - Automated surveillance
  - Digital devices
  - Sensors, actuators, transponders, meters (IoT)
  - Interactions and transactions
- Volunteered
  - Social media
  - Sousveillance/wearables
  - Crowdsourcing/neogeography
  - Citizen science



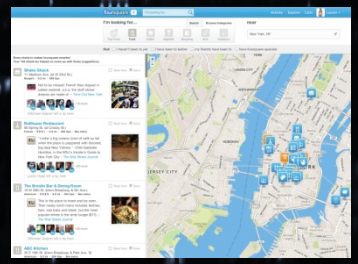



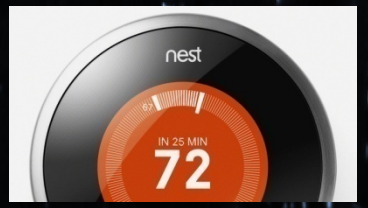


Urban big data

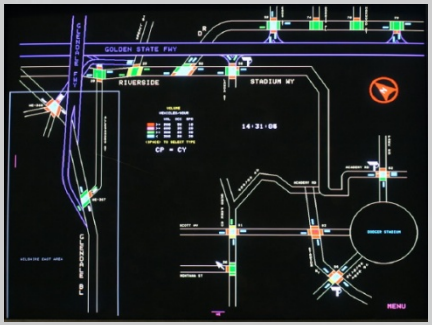
- Diverse range of public and private generation of fine-scale (uniquely indexical) data about citizens and places in real-time:
  - utilities
  - transport providers
  - environmental agencies
  - mobile phone operators
  - social media sites
  - travel and accommodation websites
  - home appliances and entertainment systems
  - financial institutions and retail chains
  - private surveillance and security firms
  - remote sensing, aerial surveying
  - emergency services
- Producing a data deluge that can be combined, analyzed, acted upon

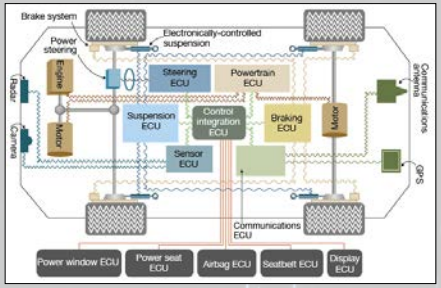


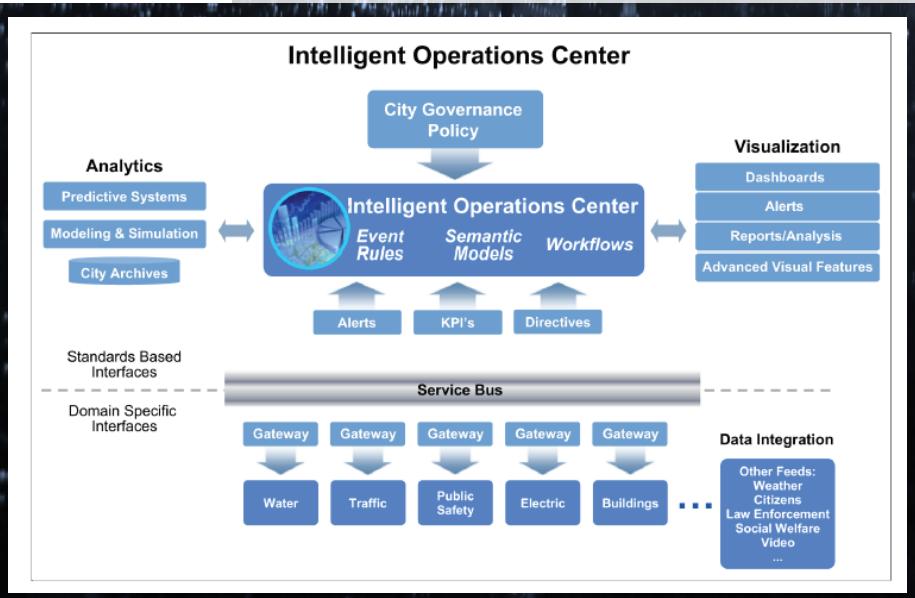


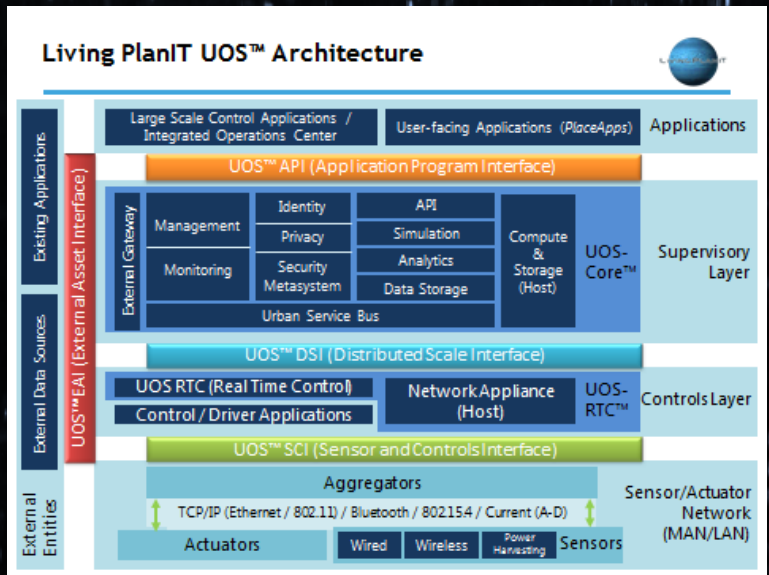
Single systems





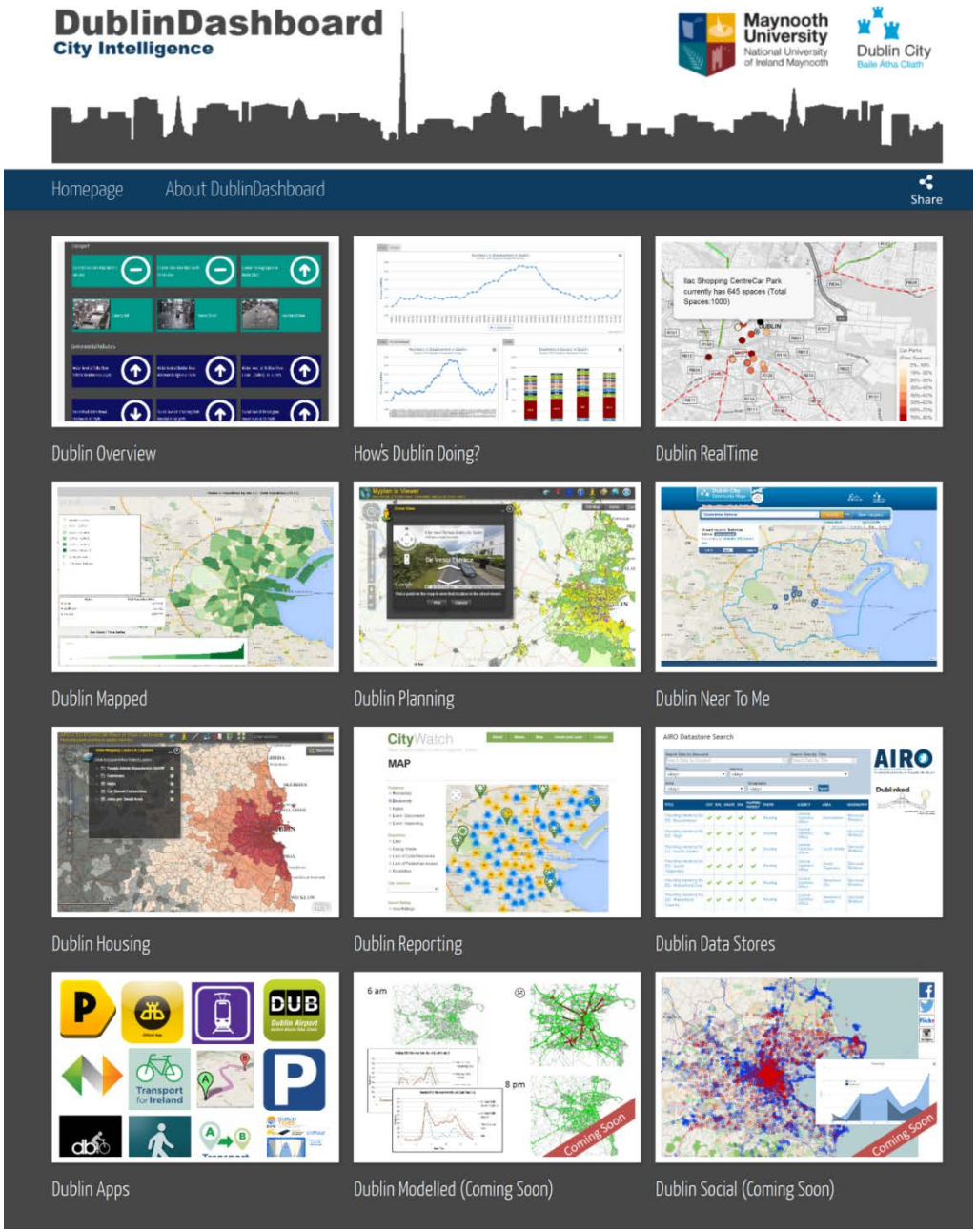
Integrated, city & sector wide

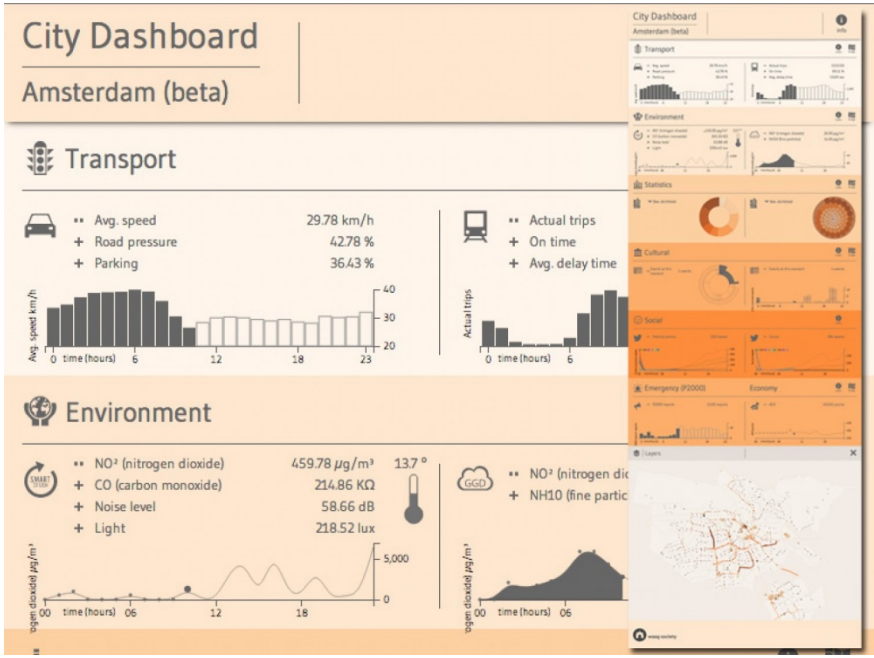


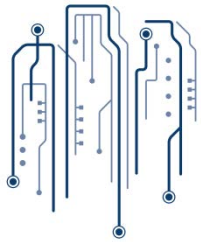


Data-driven urbanism

www.dublindashboard.ie





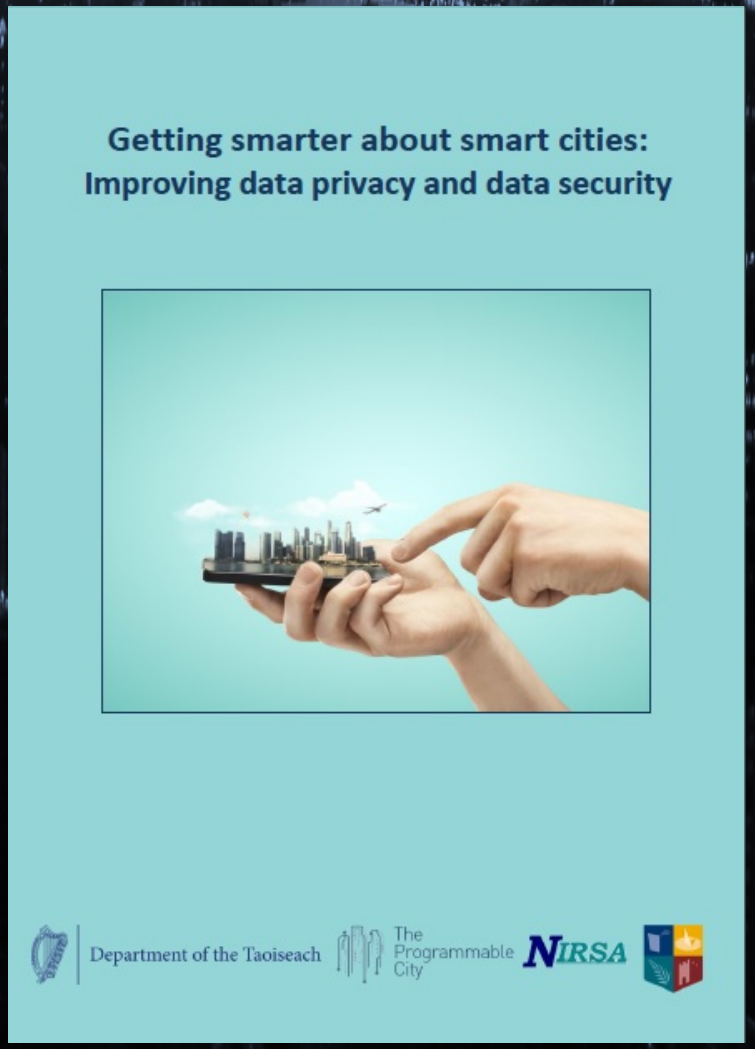


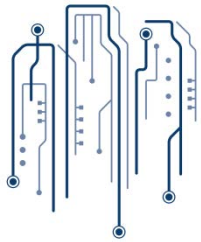
Creating smart cities

- Tackle pressing issues
- New forms of operational governance
- More efficient, competitive and productive service delivery
- Increase resilience and sustainability
- More transparency and accountability
- Enhance participation in city life and quality of life
- Stimulate creativity, innovation, entrepreneurship and economic growth
- Improve models and simulations for future development

Ethics of data-driven urbanism

- Data-driven urbanism raises all kinds of ethical & related questions
  - Data ownership and control
  - Data integration and data markets
  - Data security and integrity
  - Data protection and privacy
  - Data quality and provenance
  - Data uses
  - Data practices and politics

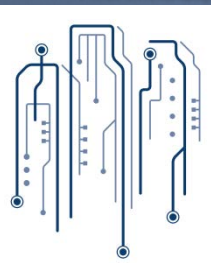




The politics of urban data

- Big data and dashboards are not simply technical tools
- Data are not pragmatic, neutral, objective, non-ideological;
- Data cannot speak for themselves
- Data are framed
- Big data and dashboards express a normative view
- And they have normative effect

The politics of urban data



Digital socio-technical assemblage

System/process
performs a task

- Reception/Operation (user/usage)
- Interface
- Code/algorithms (software)
- Data(base)
- Code Platform (operating system)
- Material Platform (infrastructure – hardware)

Context
frames the system/task

- Systems of thought
- Forms of knowledge
- Practices
- Finance
- Political economies
- Governmentalities & legalities
- Organisations and institutions
- Subjectivities and communities
- Marketplace
- Places

Privacy and big urban data

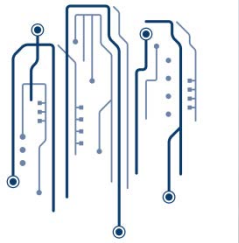
- Privacy debates concern acceptable practices with regards to accessing and disclosing personal and sensitive information about a person
  - *identity privacy* (to protect personal and confidential data)
  - *bodily privacy* (to protect the integrity of the physical person);
  - *territorial privacy* (to protect personal space, objects and property);
  - *locational and movement privacy* (to protect against the tracking of spatial behaviour)
  - communications privacy (to protect against the surveillance of conversations and correspondence);
  - transactions privacy (to protect against monitoring of queries/searches, purchases, and other exchanges)

A Taxonomy of Privacy Harms (compiled from Solove 2006)

Domain	Privacy breach	Description
Information Collection	*Surveillance*	Watching, listening to, or recording of an individual's activities
	Interrogation	Various forms of questioning or probing for information
Information Processing	*Aggregation*	The combination of various pieces of data about a person
	Identification	Linking information to particular individuals
	Insecurity	Carelessness in protecting stored information from leaks and improper access
	Secondary Use	Use of information collected for one purpose for a different purpose without the data subject's consent
	Exclusion	Failure to allow the data subject to know about the data that others have about her and participate in its handling and use, including being barred from being able to access and correct errors
Information Dissemination	*Breach of Confidentiality*	Breaking a promise to keep a person's information confidential
	Disclosure	Revelation of information about a person that impacts the way others judge her character
	Exposure	Revealing another's nudity, grief, or bodily functions
	Increased Accessibility	Amplifying the accessibility of information
	Blackmail	Threat to disclose personal information
	Appropriation	The use of the data subject's identity to serve the aims and interests of another
	Distortion	Dissemination of false or misleading information about individuals
Invasion	*Intrusion*	Invasive acts that disturb one's tranquillity or solitude
	Decisional Interference	Incursion into the data subject's decisions regarding her private affairs

Privacy and big urban data

- Intensifies datafication



Location/movement data

- *Controllable digital CCTV cameras* + ANPR + facial recognition
- *Smart phones*: cell masts, GPS, wifi
- *Sensor networks*: capture and track phone identifiers such as MAC addresses
- *Wifi mesh*: capture & track phones with wifi turned on
- *Smart card tracking*: barcodes/RFID chips (buildings & public transport)
- *Vehicle tracking:* unique ID transponders for automated road tolls & car parking
- *Other staging points*: ATMs, credit card use, metadata tagging
- *Electronic tagging; shared calenders*

Data type	Data permissions that can be sought by android apps (from Hein 2014)
Accounts log	email log
App Activity	name, package name, process number of activity, processed id
App Data Usage	Cache size, code size, data size, name, package name
App Install	installed at, name, package name, unknown sources enabled, version code, version name
Battery	health, level, plugged, present, scale, status, technology, temperature, voltage
Device Info	board, brand, build version, cell number, device, device type, display, fingerprint, IP, MAC address, manufacturer, model, OS platform, product, SDK code, total disk space, unknown sources enabled
GPS	accuracy, altitude, latitude, longitude, provider, speed
MMS	from number, MMS at, MMS type, service number, to number
NetData	bytes received, bytes sent, connection type, interface type
PhoneCall	call duration, called at, from number, phone call type, to number
SMS	from number, service number, SMS at, SMS type, to number
TelephonyInfo	cell tower ID, cell tower latitude, cell tower longitude, IMEI, ISO country code, local area code, MEID, mobile country code, mobile network code, network name, network type, phone type, SIM serial number, SIM state, subscriber ID
WifiConnection	BSSID, IP, linkspeed, MAC addr, network ID, RSSI, SSID
WifiNeighbors	BSSID, capabilities, frequency, level, SSID
Root Check	root status code, root status reason code, root version, sig file version
Malware Info	algorithm confidence, app list, found malware, malware SDK version, package list, reason code, service list, sigfile version

Privacy and big urban data

- Intensifies datafication
- Deepens inferencing
- Weak anonymization and enables re-identification
- Opacity and automation creates obfuscation and reduces control
- Data are being shared and repurposed and used in unpredictable and unexpected ways
- Notice and consent is an empty exercise or absent
-

Fair Information Practice Principles (OECD, 1980)	
Principle	**Description**
Notice	Individuals are informed that data are being generated and the purpose to which the data will be put
Choice	Individuals have the choice to opt-in or opt-out as to whether and how their data will be used or disclosed
Consent	Data are only generated and disclosed with the consent of individuals
Security	Data are protected from loss, misuse, unauthorized access, disclosure, alteration and destruction
Integrity	Data are reliable, accurate, complete and current
Access	Individuals can access, check and verify data about themselves
Use	Data are only used for the purpose for which they are generated and individuals are informed of each change of purpose
Accountability	The data holder is accountable for ensuring the above principles and has mechanisms in place to assure compliance

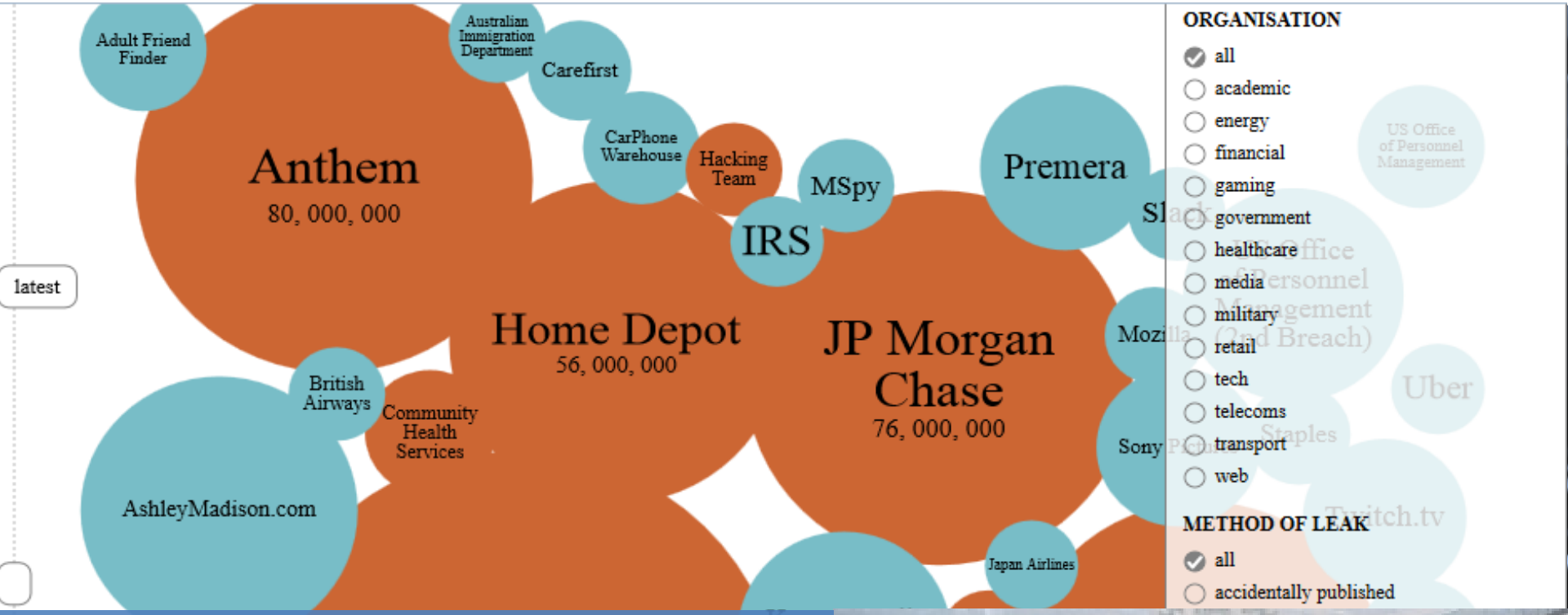
Redundant in the age of big urban data?

Privacy and big urban data

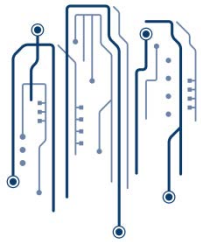
- Intensifies datafication
- Deepens inferencing
- Weak anonymization and enables re-identification
- Opacity and automation creates obfuscation and reduces control
- Data are being shared and repurposed and used in unpredictable and unexpected ways
- Notice and consent is an empty exercise or absent
- Hacking cities and data security

http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

Hacking the City?

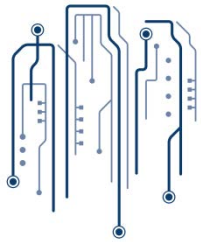


- Weak security and encryption
- Insecure legacy systems and poor maintenance
- Large and complex attack surfaces and interdependencies
- Cascade effects
- Human error and disgruntled (ex)employees



Suggested solutions

- Market:
  - Industry standards and self-regulation
  - Privacy/security as competitive advantage
- Technological
  - End-to-end strong encryption, access controls, security controls, audit trails, backups, up-to-date patching, etc.
  - Privacy enhancement tools
- Policy and regulation
  - FIPPs
  - Privacy by design;
  - security by design
- Governance
  - *Vision and strategy*: (1) smart city advisory board and smart city strategy;
  - *Oversight of delivery and compliance*: (2) smart city governance, risk and compliance board;
  - *Day-to-day delivery:* (3) core privacy/security team, smart city privacy/security assessments, and (4) computer emergency response team

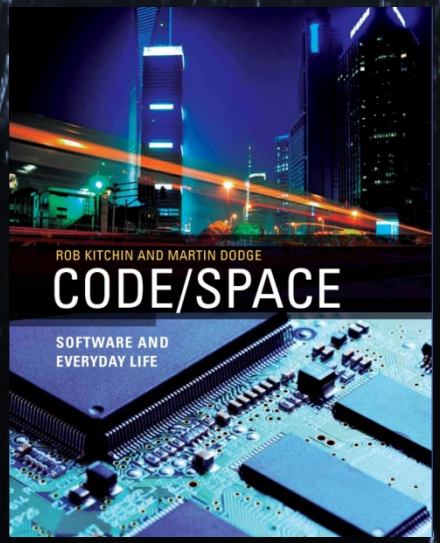


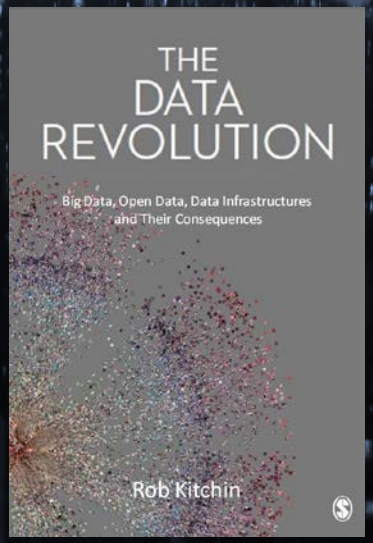
Conclusion

- Entering an era of embedded and mobile computation
- Vast quantities of real-time data, cities are responsive to these data, and enable new kinds of monitoring, regulation and control
- Cities are becoming data-driven and are enacting new forms of algorithmic governance
- Whilst data-driven, networked urbanism undoubtedly provides a set of solutions for urban problems it also raises a number of ethical and normative questions
- The challenge is to realise the benefits whilst minimizing pernicious effects

Rob.Kitchin@nuim.ie
@robkitchin

http://www.nuim.ie/progcity
@progcity





The Programmable City

Maynooth University
National University of Ireland Maynooth

AIRO
All-Island Research Observatory
An tIonad Breathnaithe um Thaighde Uile-Éireann

erc European Research Council

SFI Science Foundation Ireland