

Innovating in the Digital Ecology: Social Issues and Consequences

Professor Robin Mansell

London School of Economics and Political Science

**HEPTech Academia Meets Industry on Big Data ICT1,
Budapest 30-31 March 2015**



MEDIA@LSE
Department of Media and Communications

What we know – ‘big data’

- Internet carries communications of 2.4 billion internet users.
- In one minute those 2.4 billion transfer 1,572,877 gigabytes of data, including 204 million emails.
- 4.1 million Google searches, 6.9 million messages sent via Facebook, 347,222 posts to Twitter and 138,889 hours of video watched on YouTube.

(“What happens in an internet minute?” Intel Corp, 5 Dec 14)



Deluge!!!

Data!!

Security

Marketing

Scientists

Social Scientists

Humanists

Funding agencies

Policy makers

Librarians

Internet architects

Publishers

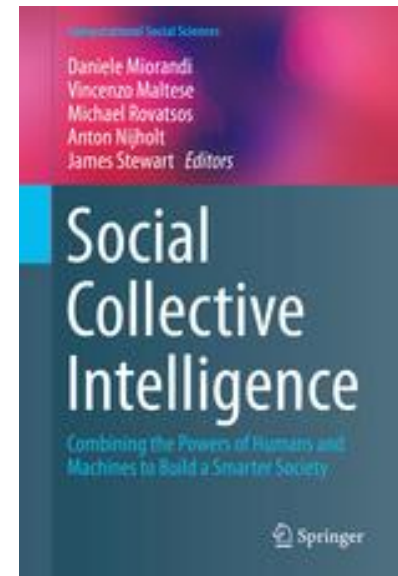
Adapted: Prof David de Roure, Oxford University



What we don't know: Collective Intelligence

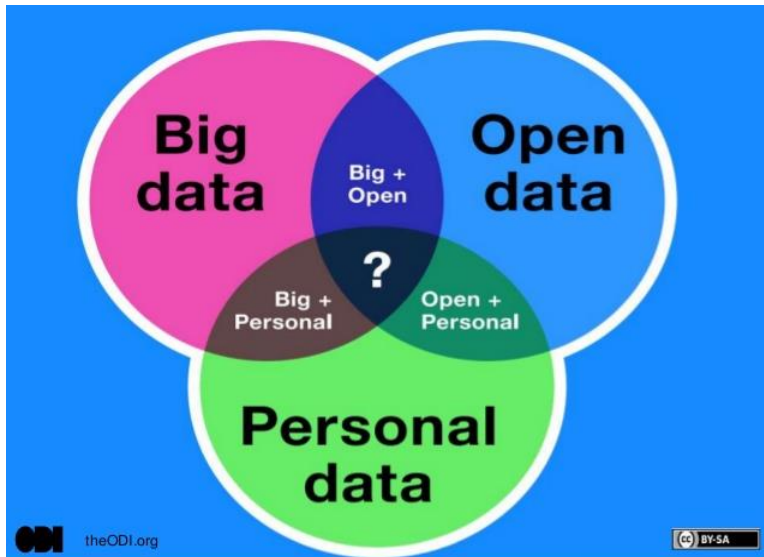
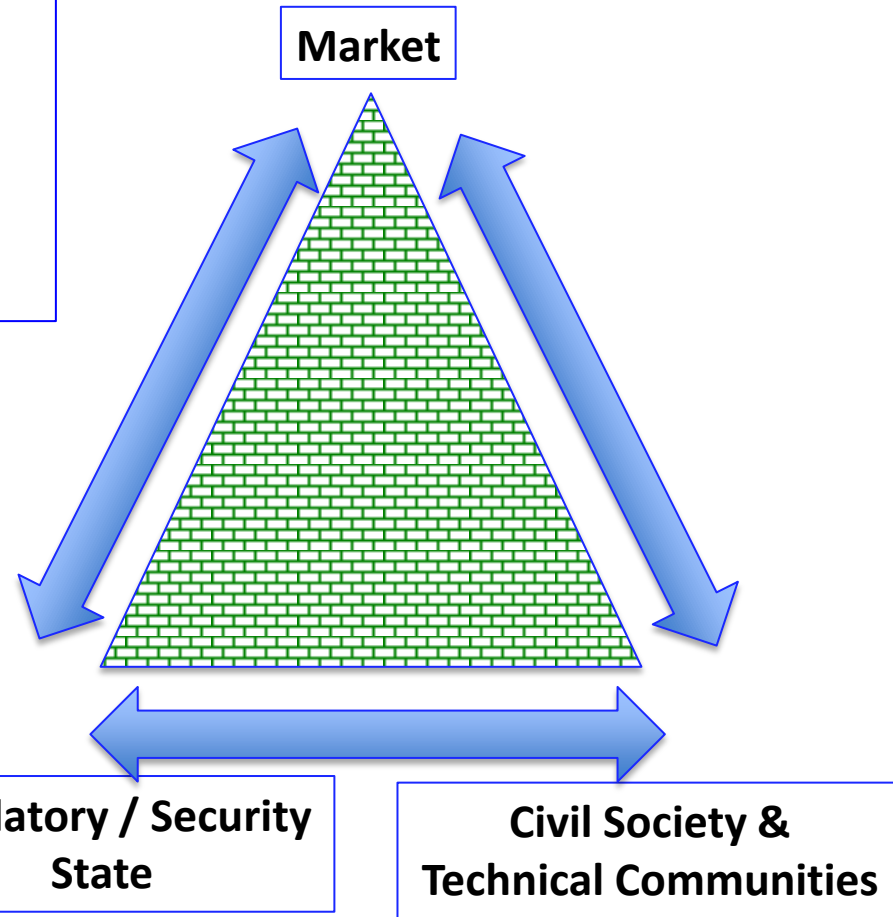
How to:

- give technological and social aspects *equal* weight.
- develop incentives, mechanisms and decision – making algorithms that can drive desirable system-level behaviour
- exploit collective experience through intentional design and *co-evolving governance* structures and processes.

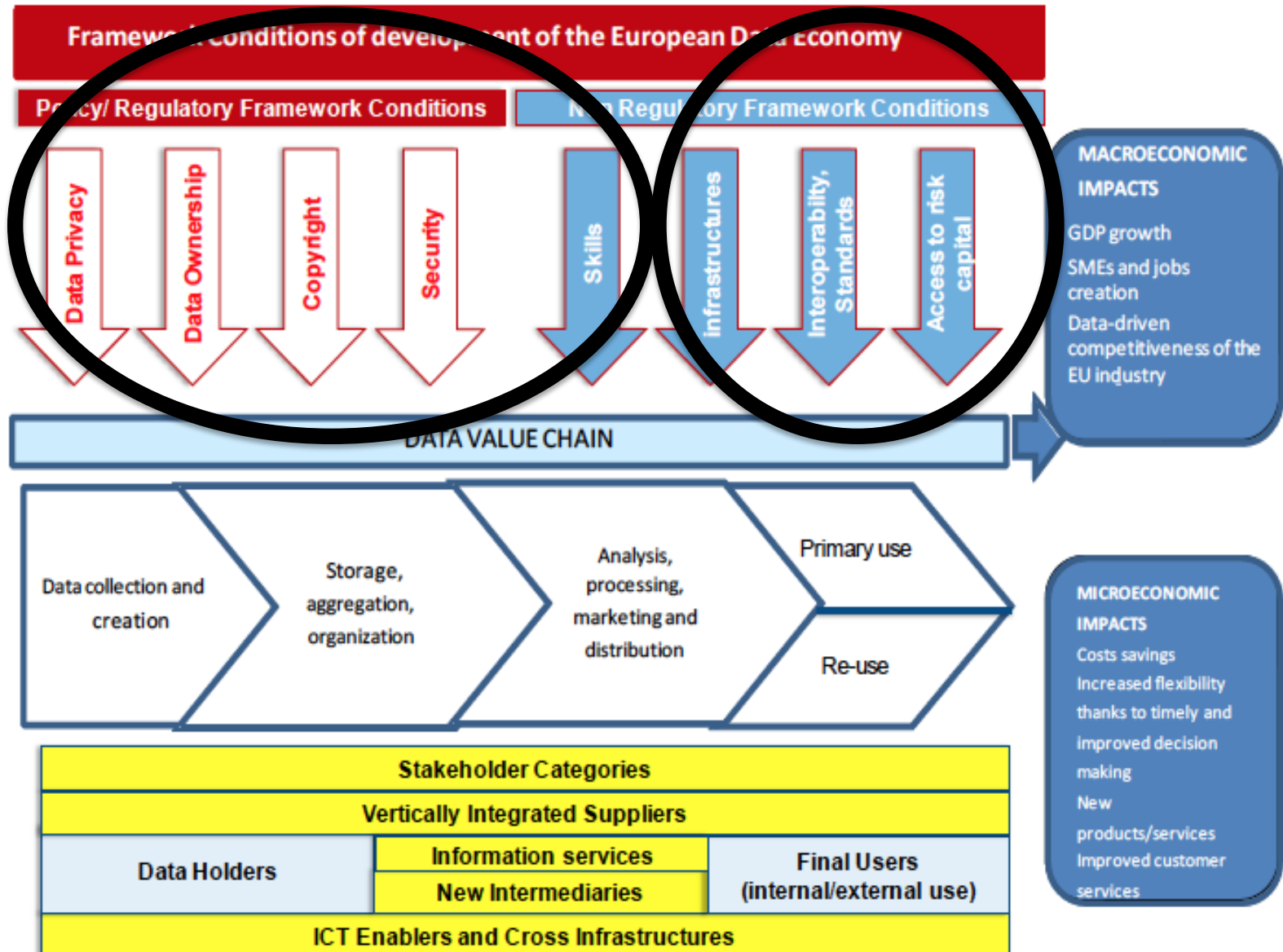


Social Science -

Who has (should have) authority to govern 'big data' activities?



What we don't know – 'Big Data'



Governance – Conflicts & Controversies

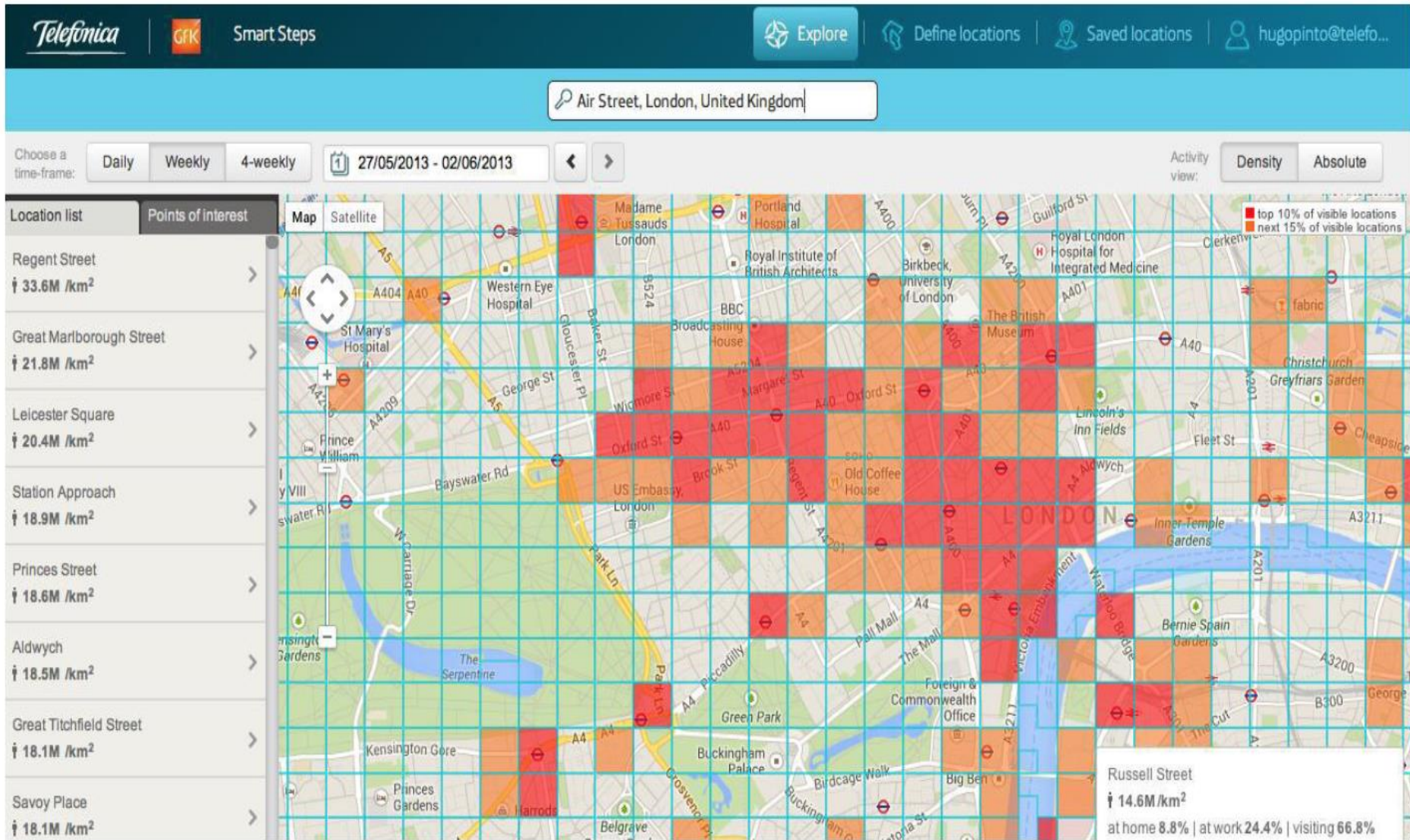
- Privacy
- Data Ownership (control)
- Copyright/IPRs
- Security/Surveillance



- Hierarchical and / or polycentric decision making.
- Balancing conflicting values - improvised action by algorithms & humans.
- Achieving the 'right amount of regulation'.

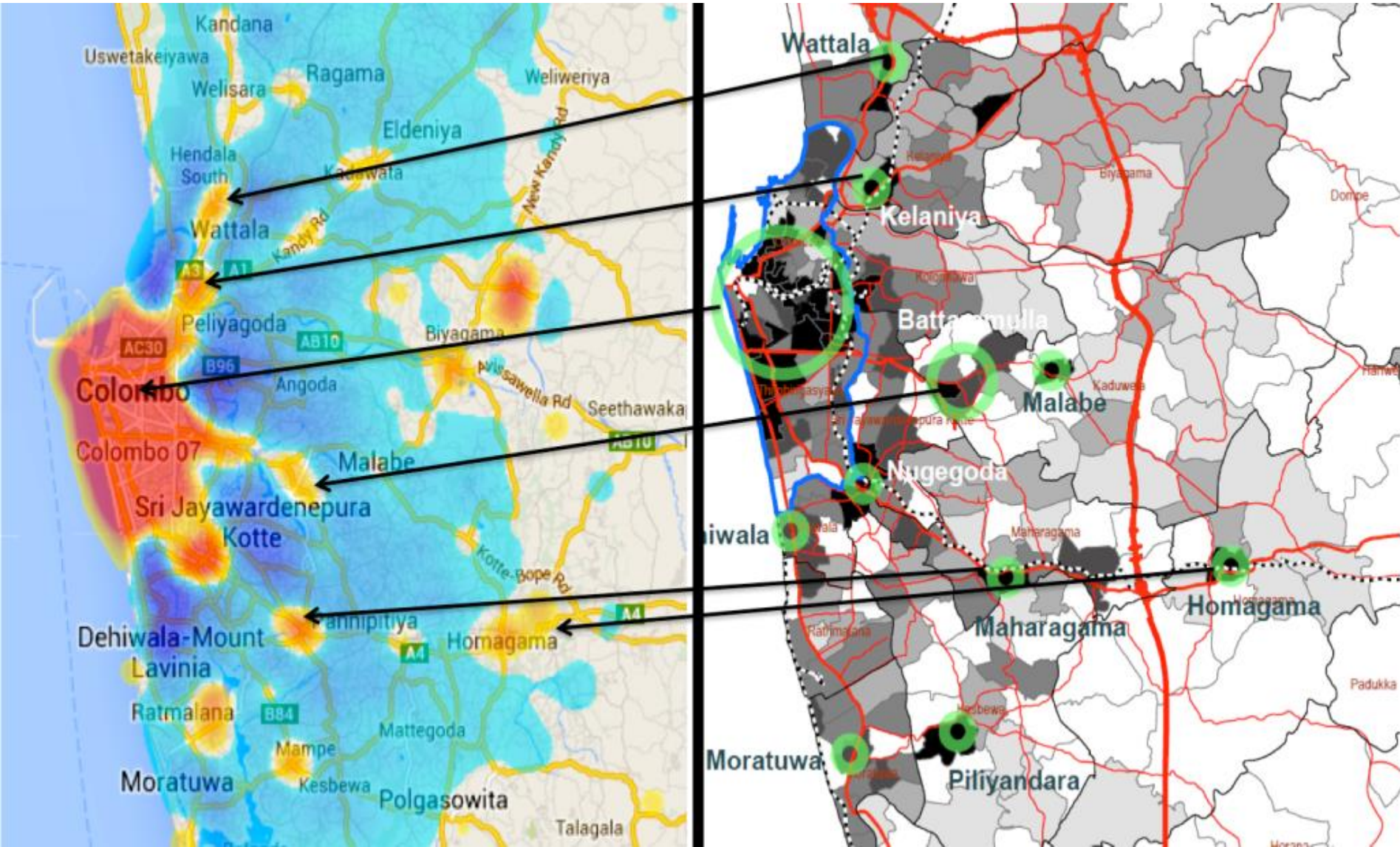
Example 1: Retail & 'Big Data'

'Smart Steps' – Telefonica - Morrisons



Example 2: Transport & 'Big Data'

Big (Mobile) Data, Colombo, Sri Lanka



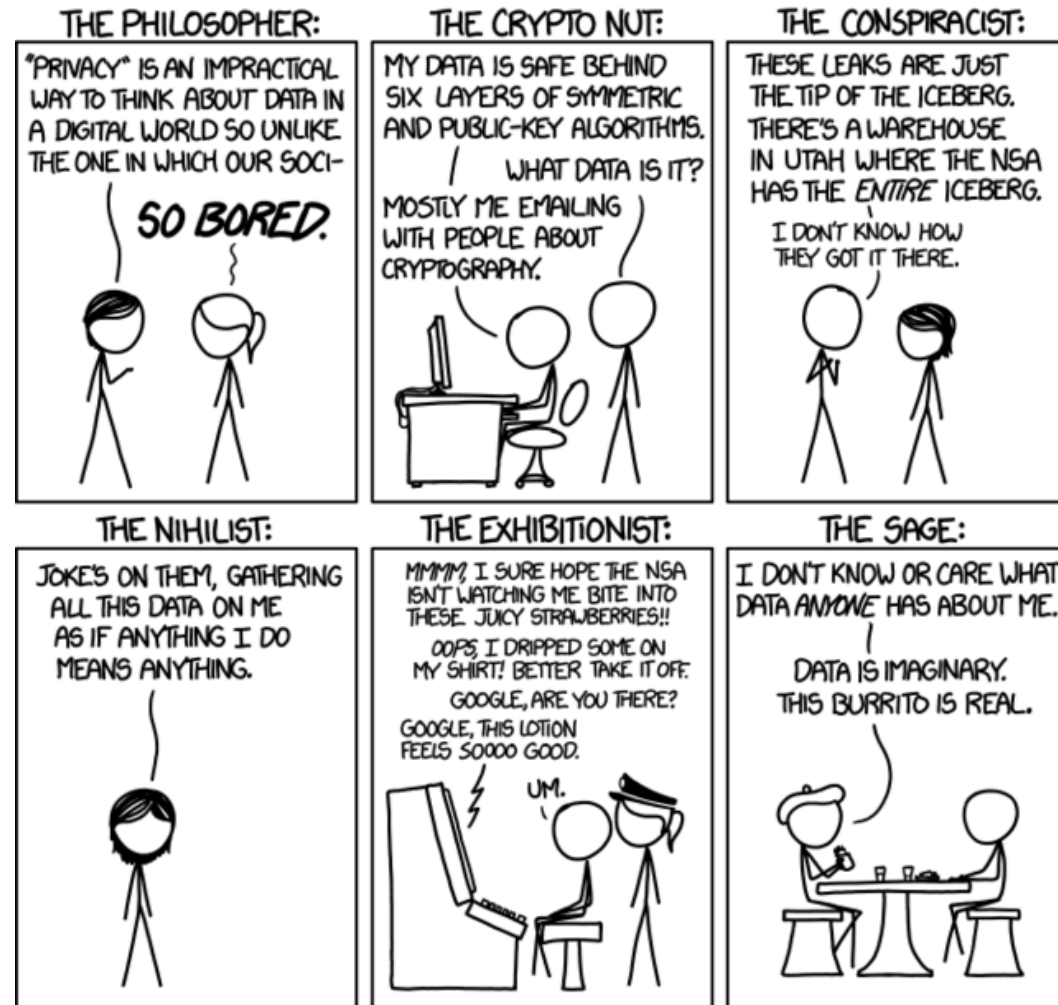
Governance 1: Privacy & Data Protection

EU Data Protection Directive

Explicit consent, transparency and fairness, accuracy, data portability, right to object to use of data for marketing, right of erasure, right not to be subject to automated processing decisions.

Aim: traceable anonymity or complete anonymity?

OPINIONS ON INTERNET PRIVACY



Governance 2: Security / Surveillance

“Some rights are not absolute ... which means that there may be circumstances in which it is appropriate to interfere”.

- Treatment of Data
Communications Plus - web browsing history
- Concerns about use of encryption ‘by default’
- Increasing commercial value of data traces

(Intelligence and Security Committee of Parliament UK, March 2015).

Resistance to data analytics by civil liberties activists & critical social science community:

Sousveillance but against predictive data mining and social sorting/targeting.

'Big Data' - Uneven Sector Development

Data about 'Big Data' are weak, few cases in public sector – barriers are technical, organizational & governance related.

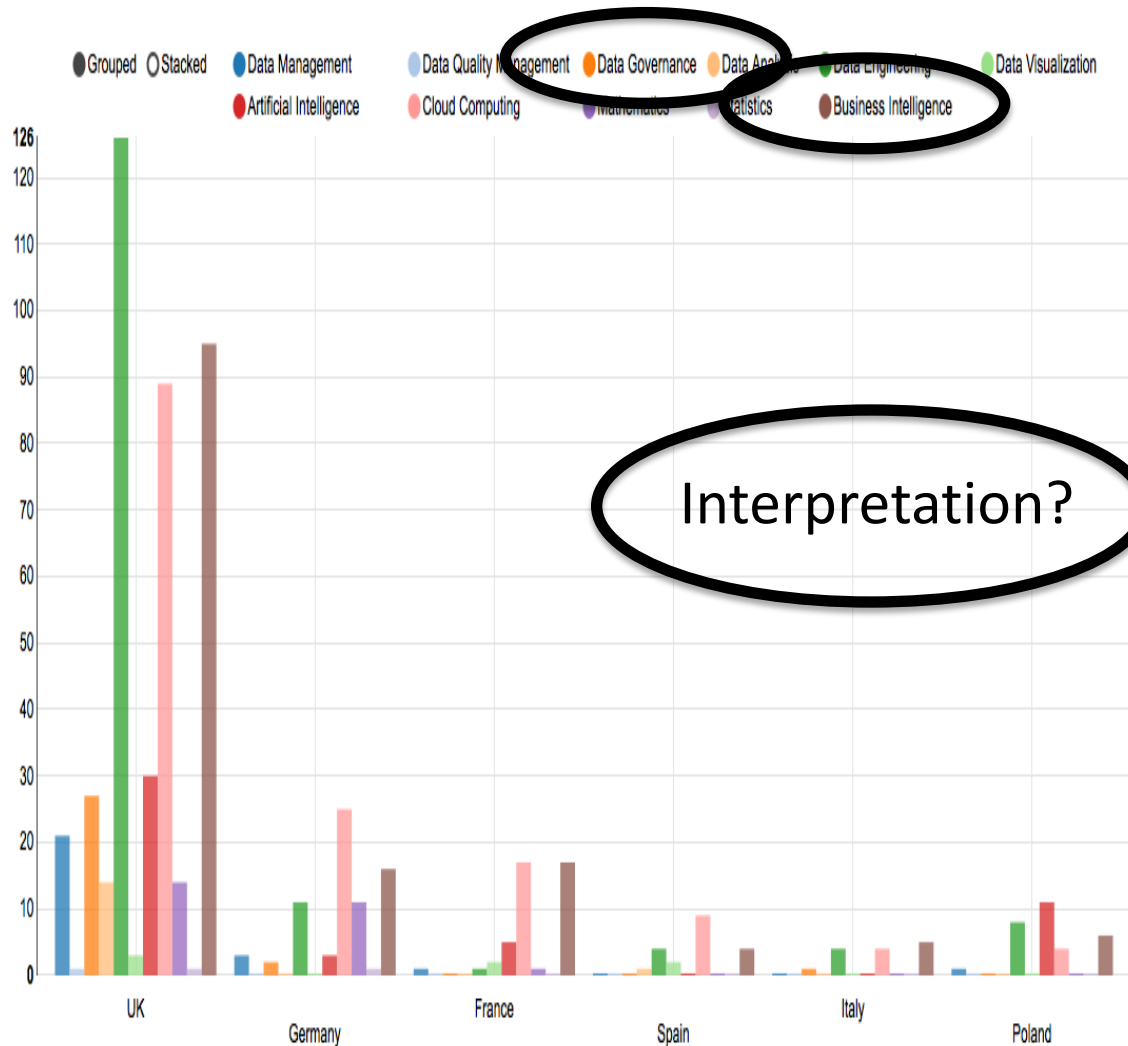
Western Europe	Volume	Variety	Velocity	Value	Intensity of Big Data Drivers
Finance	High	Medium	Hot	Hot	High
Process Manufacturing	Low	Low	Low	Medium	Low
Discrete Manufacturing	Low	Low	Medium	High	Medium
Retail/Wholesale	High	Hot	Hot	High	Hot
Telecom/Media	Hot	High	Hot	Hot	Hot
Utilities/Oil & Gas	High	High	Hot	Hot	Hot
Prof. Services/Transport	Medium	Low	Medium	Medium	Medium
Government/Education	Low	Low	Medium	Medium	Medium
Healthcare	Medium	Medium	Medium	Medium	Medium
Total	Medium	Medium	High	High	High



- Hot
- High
- Medium
- Low

'Big Data' Skills Gap

- Big Data technology & services market in Western Europe to grow from USD 2.3 bn (2013) to 6.8 bn in 2018; CAGR of 24.6% (IDC Europe).
- By 2018 demand for data-competent managers and analysts in US will be 450,000; supply will be short at 160,000 (McKinsey 2014).



Barriers to 'Big Data' Development

- **Accessing data** – companies reluctant to share, government open data initiatives cover limited sectors.
- **Privacy** – rights-based approaches are not global; 'informed consent' models increasingly impractical; policies do not deal with secondary use of data; unclear boundary between personal and non-personal data.
- **Analytical/Interpretation** – 'rubbish in, rubbish out' – data quality and provenance, measurement bias (correlation is not causation) and need for experimental techniques.

Conclusion: Imagining 'Big Data' Futures

Computational Justice –

Is it workable? Studies intentionality and justice.

Application in multi-agent models. Self-governance through formal representation of data access, copyright, and privacy norms, all embedded in rule- based models.

“To be effective, a data analyst needs to turn data into information, information into knowledge, and knowledge into action. You can't do this without communication”.

Need for “serious benefit-cost analysis to guide regulatory policy”.

(Hal Varian, Google, April 2015)

