

Open/Free Data, Big Data, 'Official' Data and Crowdsourced Data

Some basic questions:

- Is integration necessary (or just desirable)?
- Is integration possible and practicable – within budgets and human resources available?
- How far away are we from achieving such integration?
- Where do the challenges lie?

What is Open Data?

Open Definition from the Open Knowledge Foundation:

- ❖ principles that define “openness” in relation to data and content,
- ❖ precisely defines “open” in the terms “open data” and “open content”,
- ❖ ensures interoperability (shared access) between different collections of open material.

<http://opendefinition.org/okd/>

<http://okfn.org/>

What is Open Data?

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike.”

<http://opendefinition.org/od/> <http://okfn.org/>



OKN's Open Data Definition

The Open Knowledge Foundation's definition covers:

- Access
- Redistribution
- Reuse
- Absence of Technological Restriction
- Attribution
- Integrity
- No Discrimination Against Persons or Groups
- Distribution of License
- License Must Not Be Specific to a Package
- License Must Not Restrict the Distribution of Other Works

<http://opendefinition.org/od/>

Open Data Census

Global Census Facts at 2015

- ❖ Number of countries = 122
- ❖ Number of datasets = 1586
- ❖ Number of open datasets = 156
- ❖ **Percentage open = 9%**



From the Open Data Index
<https://index.okfn.org/>

G8 Open Data Charter

- Principle 1 – Open Data by default
- Principle 2: Quality and Quantity
- Principle 3: Usable by All
- Principle 4: Releasing Data for Improved Governance
- Principle 5: Releasing Data for Innovation



Open Data Challenges

1. What data should be made public?
2. How to make data publicly 'open'?
3. How to efficiently implement and monitor Open Data policy?
 - Regulatory/Statutory reporting requirement?
4. How to judge the effectiveness of an Open Data policy?
 - Indicators?



Making Data Publicly 'Open'

1. **Agreeing data (& service) standards**
 - ... and introducing them.
2. **Setting appropriate policies**
 - ... and enacting them.
3. **Promulgating regulations**
 - ... and enforcing them.



Lessons learned from the EU's PSI Re-use Directive(s) (2003 and 2013)

Monitoring Policy

1. Monitoring Open Data

- Should you monitor Open Data policy
- Can you monitor Open Data policy?

2. Implementing policy

- Voluntary v. mandatory
- Regulations?
- Handling infringements

3. Technology

- For monitoring and reporting



Judging Effectiveness

1. How to judge the effectiveness of a government's Open Data policy?
2. Defining 'effectiveness'
 - Benefits for government, society and businesses
 - Cost-Benefit Analysis – feasible?
 - Identifying tangible v. intangible benefits
3. What 'indicators of success' to use
 - Some will be financial
 - Many (intangibles) will be difficult to measure



Crowdsourcing – Impact and Challenges

- **Crowdsourcing consists in outsourcing tasks to a broad external group of people which contributes to an overabundant public dataset.**
- **Data arise from a network of heterogeneous devices, sensors or sources, and involve the participation of various applicants (experts and novices).**
- **Reputation models have been precisely designed for trust management.**
- **Participatory sensing places individuals in their environment and neighborhood community via cloud services that collect and analyze systematic data.**
- **Modern smartphone functionalities provide a convenient medium for collecting user feedback in addition to raw data collected by the sensors on the device.**

Crowdsourcing – Impact and Challenges

- **Crowdsourcing provides a huge volume of data and valuable feedback over a wide coverage area.**
- **Google Live Traffic to Google Maps which exploits the position data sent by smartphones with Android operating system allowing real-time traffic information**
- **Free GPS application Waze aims at generating traffic information in real-time - available for Android, IOS, Windows Mobile, Symbian and BlackBerry (2.5 million French users in February 2013, 70 million users in the world in February 2014). The application allows the users to add new roads, to report accidents, traffic jams and speed traps.**
- **Road surface condition monitoring can be achieved through accelerometers embedded in smartphones, as addressed by the TrafficSense project led by Microsoft.**

Crowdsourcing – Impact and Challenges

- **Environmental conditions can be monitored via devoted smartphone applications.**
 - **The aBluSen application is a Bluetooth-based temperature and humidity acquisition system.**
 - **Air quality monitoring by means of smartphones includes the GasMobile architecture that relies on hardware consisting of a low-cost ozone sensor directly connected to the smartphone through USB host mode and an Android API.**
- **Global Eye on Earth initiative and programme depends heavily on crowdsourced environmental information as one of its founding principles.**