Where are we going?

Delivering open data in Ireland and Europe

Eoin MacCuirc

Databank and Dissemination

Central Statistics Office, Ireland

eoin.mccuirc@cso.ie
"I liked it better before big data and metadata when we just had good old regular data."
Outline

• Open data, the value of official statistics, the data deluge?
• Data silos, opening data, open data licenses, open data formats
• CSO journey from paper to SPARQL on a shoestring
• CSO, Ireland and the Open Government Partnership publishing open data the context, strategy and resources
• What’s happening in Europe?
• Lessons learned
The challenge of building more open, transparent and accountable public governance in Ireland is one of my key objectives as Minister; and this is being achieved in terms of the range of reform initiatives currently being implemented across public bodies. The Open Data Initiative is a key part of these reform activities. It is closely aligned with the Public Service Reform Plan, the ICT Strategy, and the Civil Service Renewal Plan. It is also a core element of Ireland’s Open Government Partnership (OGP) National Action Plan which I launched in July 2014. **Open Data is a particular interest of mine. I see it as the new basic resource of the 21st century; and publishing it for re-use offers the possibility of real economic, social and democratic benefits.**

**Speech by Minister for Public Expenditure and Reform, Mr Brendan Howlin TD at Open Data Event 30 June 2015**
While the data deluge accelerates, most of the data produced remains locked in deep Web databases. For the linked open data to benefit from the potential represented by this huge amount of data, it is crucial to come up with solutions to expose heterogeneous databases as linked data.

Franck Michel, Loïc Djimenou, Catherine Faron-Zucker, and Johan Montagnat
Government Data shall be considered “open” if the data is made public in a way that complies with the principles below:

1. Data must be **complete**. All public data is made available.
2. Data must be **primary**. Data is published as collected at the source, with the finest possible level of granularity, and not in aggregate or modified forms.
3. Data must be **timely**. Data is made available as quickly as necessary to preserve the value of the data.
4. Data must be **accessible**. Data is available to the widest range of users for the widest range of purposes.
5. Data must be **machine-processable**. Data is structured so that it can be processed in an automated way.
6. Access must be **non-discriminatory**. Data is available to anyone, with no registration requirement.
7. Data formats must be **non-proprietary**. Data is available in a format over which no entity has exclusive control.
8. Data must be **license-free**. Data is not subject to any copyright, patent, trademark or trade secrets regulation. Reasonable privacy, security and privilege restrictions may be allowed as governed by other statutes.
9. **Permanence** refers to the capability of finding information over time.
10. **Usage costs** One of the greatest barriers to access to ostensibly publicly available information is the cost imposed on the public for access – even when the cost is de minimus.

Florian Bauer, Martin Kaltenböck
Ideally

- Associated with an open data license
- Published in open data format(s)
- Machine readable, streamable, bulk downloadable, machine discoverable
- Software and hardware agnostic
- Non-proprietary
- Using standardised metadata
- Linkable
CSO Open Data - 5 star data target

- PDF – one star
- Electronic release - one star (responsive design)
- StatBank – three star (full file download)
- JSON API – three star
- data.cso.ie – five star
- Open Cube project – five star
- OSI collaboration – five star
CSO Open Data Journey

I am learning to trust the journey, even if I don't understand it.

POWEROFPOSITIVITY.COM
CSO Guiding Principle

Single source multiple output
Moving away from Paper
CSO was one of four international pilot sites for the EU Open Cube project for publishing and enriching Linked Open Statistical Data. CSO is working with INSIGHT NUIG https://www.insight-centre.org/ experts in Linked Data and the Semantic Web for our data. The aim of the CSO is to publish all CSO time series data on StatBank and the Public Sector Statistics Network in RDF format. Again data in the RDF format is linkable on the semantic web adding to its openness (**** and *****)
CSO Linked open data publication

• Get the data on StatBank & PSSN
• Convert the data to JSON-Stat
• Agree a national URI strategy
• Convert the JSON-Stat to RDF
• Execute some pathfinder projects to build linkages
• Collaborate with others locally, nationally and internationally
Open Government Partnership

Welcome to the Open Government Partnership Ireland website. The Open Government Partnership (OGP) is an initiative of 70 countries. It aims to secure concrete commitments from governments to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance. It is run by a steering committee composed of governments and civil society organisations.

Ireland's Open Government Partnership
National Action Plan 2016-2018

On Wednesday 7th December the Irish Government published Ireland’s Open Government Partnership National Action Plan 2016-2018. This is available in PDF and Word formats. Shortly, it will also be published in HTML format.

A table setting out the proposed commitments and the response of the Department of Public Expenditure and Reform to those is also available.

Creating a new
National Action Plan
2016-2018
Operational Strategy

9 What are the Key Principles of this Open Data Strategy?

The Open Data Strategy will outline a coherent and consistent approach to the publication of Open Data by public bodies. Processes developed to support implementation of this Strategy, including the Technical Framework for Open Data will lead to persistent and repeatable processes that can be shared among public bodies and better facilitate the release of appropriate datasets based on real demand.

The Open Data Strategy will support the goal of the Public Service Reform Plan to achieve better outcomes for citizens, businesses and public servants. The Public Service Reform Plan envisages that these better outcomes will be delivered through a focus on service users, on efficiency and on openness, underpinned by a strong emphasis on leadership, capability and delivery.

1 The needs of citizens and businesses are at the centre of Ireland’s approach to Open Data
2 Open Data is considered the default option for appropriate new datasets. Where requested datasets are not released as Open Data, the responsible public body will provide reasons why not
3 Open Data, linked though the portal data.gov.ie, will meet the requirements of the Open Data Technical Framework over time:
   - Licence (allowing people to use data)
   - Formats (e.g., using non-proprietary instead of proprietary formats)
   - Metadata (precise descriptors about datasets)
   - Standards (to ensure a common understanding of the data)
   - Unique Resource Identifiers

Principles of Open Data Strategy
Open Government Partnership - initial three star data target

Publication of datasets via the portal will be based on the requirements of the Open Data Technical Framework, which has been developed in collaboration with the Public Bodies Working Group.

The Open Data Strategy will use the 5-star deployment scheme for Open Data. The greater the number of stars, the more reusable the data is, and the easier it is to reuse and interconnect data.

While public bodies are encouraged to maximise the openness of their formats and move towards 4 and 5 Star linked data, this Strategy recognises that this will take time and public bodies may adopt a phased approach to reaching required Open Data formats.

Action
The initial minimum target of this Strategy is for Open Data to be published at a minimum 3-Star format (non-proprietary machine-readable format).

Accordingly, all datasets published on the Open data portal will, within an agreed timeframe, be:

- Machine-Readable
- Non-Proprietary
- Associated with an Open Licence
- Described by standardised Metadata

Action
Where possible, public bodies should aim to publish as linked data, and specific targets will be set in this regard.
Open Data Irish Context

- ICT Strategy and eGovernment Strategy
- Action Plan for Jobs & “Big Data”
- EU & International Initiatives
- Open Government Partnership Action Plan
- Civil Service Renewal Plan
- Public Service Reform Plan
- Public Bodies Business Plans & Strategies
Open Data – Role in Open Government

How to make open government work?

- Citizens
- Civil society
- Private sector
- Collaboration
- Transparency
- Open Data
- Open Process
- Participation
- Open Government
- Open Service

Public data should not be locked in government offices. Here’s how Europe can benefit from #opengov #opendata #egov

@ViolaRoberto

Roberto Viola

Worldwide

Private sector

Businesses

Civil society

Social partners

Users

#opengov
Open Data Strategic Goals

1. For Public Bodies to engage in the Open Data Initiative, and ensure that publication of Open Data is embedded knowledge and information management processes.

2. For Businesses to engage with public bodies to identify datasets with commercial potential, use the datasets, and demonstrate value through use cases.

3. For Researchers to identify emerging trends for Open Data and identify future opportunities for economic, democratic and social benefits, such as Research and Development.

4. For Citizens and Civil Society to participate in the Open Data process and utilise datasets for transparency and democratic purposes.

5. The Governance Model for Open Data will identify metrics to ensure a continued evidence-based approach to future Open Data development.
Open Data Initiative Resources 2014

• **Best Practice Handbook** – this is the core document which draws together existing best practice standards for the publication and re-use of Open Data to assist in establishing best practice standards in Ireland

• **Data Audit Report** – reports on an audit of the Irish public sector datasets available online and aligns the findings with the ‘common high-value datasets’ based on the G8 Open Data Charter categories

• **Roadmap** – sets out a detailed 3 year plan for the objectives, structure and actions necessary to advance publication of national and regional Open Data for maximum impact

• **Evaluation Framework** – to assess the current state of readiness for Open Data in Ireland; to monitor and assess the ongoing progress of the initiative, and to measure the actual economic, social and political impact of the initiative.

Open Data - Technical Framework 2015

• **Technical Framework** for Open Government Data has been developed to help with the publication of official information as Open Data

• Licence

• Formats

• Metadata

• Standards

• URIs
Irish Open Data Portal
**European Open Data Portal**

**Major changes 2016**

- Impact of Open Data has increased significantly due to the publication of two studies on better service delivery and because of the high political impact of Open Data on government efficiency and effectiveness and the high social impact on increasing transparency and accountability.
- The use of Open Data has increased, partly due to the monitoring of the number of visitors on the portal.
- The presence of an Open Data Policy has increased because Ireland now has a pre-defined approach to ensure that data sets are up-to-date.

**Ireland – Conclusion**

**Print screen portal**

**Maturity 2016**

**Next steps**

- Work together with mapping institution on geospatial strategy which will contain good standards for spatial data
- Work on user interface of national portal

**Upcoming events**

- 30 Aug – 1 Sept 2016, Open Data information stand at the UN 3rd International Conference on Big Data for Statistics, in Dublin
- Open Data conference expected to take place later this year
A series of indicators have been selected to measure Open Data maturity across Europe. These indicators cover the level of development of national policies promoting Open Data, an assessment of the features made available on national data portals as well as the expected impact of Open Data.
Lessons Learned

• To publish, link and benefit from linked open data CSO needed to collaborate. Get great partners.
• There is a significant learning curve and still a lot of unknowns in publishing and linking linked open data
• It is important to develop and share common vocabularies, metadata, processes, skillsets and tools for linked open data in the ESS
• A single source of truth. Single source, multiple outputs
• Everything is connected, connect everything
• Maximum output for minimum effort
Lessons Learned

• Get a great (government) champion
• Publishing linked open data at local, national and international levels is important and data owners are not always data experts. Help data owners publish their data. The ESS can play a critical role here.
• Publishing “quick and dirty”. “Perfection is the enemy of progress” embrace MVPs (minimum viable products). Publish great data.
• The ESS can play a key role in linking open data
• Don’t be put off by the gobbledygook
• There is a long way to go, take it one step at a time
**Last Word - First Principles**

**PRINCIPLE 1**
Open data by default: all government data will be published openly by default.

**PRINCIPLE 2**
Quality and quantity: data should be released as early as possible in its original form and fully described in clear language. This principle denotes the importance of metadata and user feedback to improve quality.

**PRINCIPLE 3**
Usable by all: data will be published in open formats for humans and machines wherever possible and will be free.

**PRINCIPLE 4**
Releasing data for improved governance: governments will share their technical expertise with each other and document their own open data initiatives.

**PRINCIPLE 5**
Releasing data for innovation: G8 governments will promote open data literacy and the provision of data in machine-readable formats.