Republic of Ireland's Open Data Strategy: Observations and Recommendations

Tracey P. Lauriault, Programmable City Project, NIRSA, National University of Ireland Maynooth, County Kildare, Republic of Ireland

The Programmable City Working Paper 3

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

Executive Summary

Working Paper 3 of the Programmable City Project is a response to the Republic of Ireland's Department of Public Expenditures and Reform (DPER) Open Data launch and the reports produced by Insight at the National University of Ireland Galway (NUIG), who were awarded the contract from a government call for tender (CfT). The Working Paper provides background context to the open data plan and critically considers governance; infrastructure; records management as well as information management and information technology (IM/IT); Legal, policy and ethical frameworks; public engagement; data curation; data dissemination and publication, and evaluation. The Paper proposes reconceptualizing open data as a function of government record keeping, information management, shared services and national spatial data infrastructures as opposed to a standalone program. By doing so, it is suggested, open data simply becomes a good governance strategy and by integrating it into broader government administration information management provides it with sustainability, especially if it becomes a normalized data dissemination strategy and a public engagement mechanism. The Working Paper also includes a number of recommendations for consideration in addition to or to complement those provided by Insight. Recommendations are as follows:

1. Good Governance

- Open data should be a natural extension of good governance strategies and not stand alone programs.
- Open data should be a key component of government information management (IM), record management, IT and national spatial data infrastructures (NSDI).
- Open data should be part of a coordinated data and information dissemination strategy, which should also include publicly funded research data, scientific data, data from the humanities, and other qualitative data.
- Focus on the production and maintenance of good quality public sector, administration, research, geospatial and scientific datasets and less on commercialization, innovation and 'high value' datasets.

- Focus on data that have societal and environmental value, and also on core/framework datasets upon which other datasets can be integrated into.

2. Open Data, Records & Information Management and Thinking more Critically about Data

- Consider open Data as a good governance strategy and as part of data and information management.
- Integrate Open Data into IM/IT, Shared Services and integrate with the NSDI
- Data infrastructures are critically important, Open Data should be considered with cloud computing, high speed internet, and hardware and software.
- Open datasets should be thought of as government records (data & information) and should be managed accordingly.
- Adopt a life-cycle and data curation approach to the management, preservation and dissemination of Open Data datasets.
- Implement the NSDI and consider the CGDI principles for the NSDI and for Open Data in Ireland
- Critically reflect on data more broadly and not just as objects at the end of an information pipeline.
- Consider evaluating the contents of an open data portal to see if these can be used to construct indicators of well-being and quality of life.

3. The DPER / Insight Roadmap and the Best Practices Handbook

3.1. Governance

- Develop an open data public interest mandate, vision and mission, and clear objectives against which performance can be evaluated.
- Reconsider the organizational structure as per the schematic in Figure 5.
- Reconsider appointments on the SIG to be expertise and skills based and less political, and that appointments be made by peers.
- Create an open data institutional entity that will operationalize the work of the ODB, SIG and Working groups and integrate these with other government programs.
- Open data officers should be appointed in all government offices
- Create temporary expert working groups to develop and implement infrastructure wide practices (see figure 5).

3.2. Legal, Policy and Ethical Framework

- Develop a data and information legal and policy framework with open data as a component of it.
- Conduct an inventory of collaborative and data sharing instruments (e.g., MOU, procurement contracts, data sharing agreements, etc.).
- Assess the outputs of the *Intellectual Property Activity in Ireland Based on Existing Data* report resulting from the RfT in the spring of 2014.
- Conduct an inventory of all laws, regulation, policies and directives that would govern how data are collected and disseminated.

- Develop a set of explicit legal, policy and ethical guidelines for the management of public sector data and open data based on laws, regulation, directives, policies and practices in Ireland for public sector officials.
- Include these guidelines as part of the data dissemination decision-making tree (Figure 6).

3.3. Public Engagement

- Engage with stakeholders on developing the mission, vision and mandate for the Open Data strategy.
- Engage with stakeholders to shape how an Open Data roadmap and strategy could look.
- Engage with, study, build upon and harmonize the Open Data strategy with existing public sector data dissemination programs.
- Review and assess existing technologically mediated engagement tools and social media applications in other jurisdictions.
- Public sector officials and departments should develop processes and be receptive to evidence based public input into public policy and planning, and learn to solicit feedback from the public in a useful and educated way.
- Consider crowdsourcing, VGI and citizen science as a public engagement strategy.

4. Data Curation or a Data Audit?

- Adopt a digital data curation and life-cycle approach to the management of data and conduct the data audit accordingly.
- Adopt the Data Audit Framework.
- Ensure that additional elements are added into the data audit (e.g. geocoded, scale, time).
- The high value approach to the selection of data should be reconsidered, and an evaluation of what current data 'clients' value, should be considered.
- Recognize the limitations of a machine only audit, and broaden search criteria to include all data not just those in open formats and under an open licence.
- Conduct a full inventory of portals and catalogues from all sectors in Ireland and integrate their metadata to ensure cross disciplinary discoverability.
- Publish the results of the data audit.

5. Data Dissemination and Publication

- It is highly recommended that DPER consider adopting the well established data curation life cycle management approach similar to the one developed by the Digital Curation Centre, and consider taking a data curatorial approach in lieu of a data audit.
- Adopt the Data Audit Framework for data curation as well as those developed by the Digital Curation Centre and consider developing an Information Management Directive which incorporates the ideals of Open Data, preservation and archives.
- Create a decision making tree to help public officials determine what can and cannot be published. Figure 6 is an example to guide decisions on the management and dissemination of sensitive data.
- The outcomes of the decision derived from the application of the open data publication decision making tree would then form the basis for the decision supporting why some datasets are not published by default.

- A data management and dissemination WG should be created along with those in Figure 5, and invite experts from the Digital Repository of Ireland, library and archives and information studies, geospatial community to help develop a comprehensive access, dissemination, data management and preservation plan for Ireland.

6. Evaluation

- Assess current performance and evaluation frameworks within the Irish public sector, including auditing frameworks, or those commonly adopted and reported on in other countries that have well established Open Data programs such as Canada, the US and the UK and as per the RfT.
- Reassess the Open Data Barometer evaluation recommendation in the DPER/Insight report in light of its objectives and its target use and determine if it is a suitable model for a western developed national Open Data program.
- Consider high impact datasets, those of public, social and environmental significance along with those considered to be of high value

Introduction - Context

The Republic of Ireland's Department of Public Expenditures and Reform (DPER) launched its first Open Data Portal data.gov.ie on July 22nd (see Figure 1). It also conducted three public consultations, one with public sector bodies on July 30, and later on September 8th with civil society organizations and the public, while also accepting responses and recommendations by email.

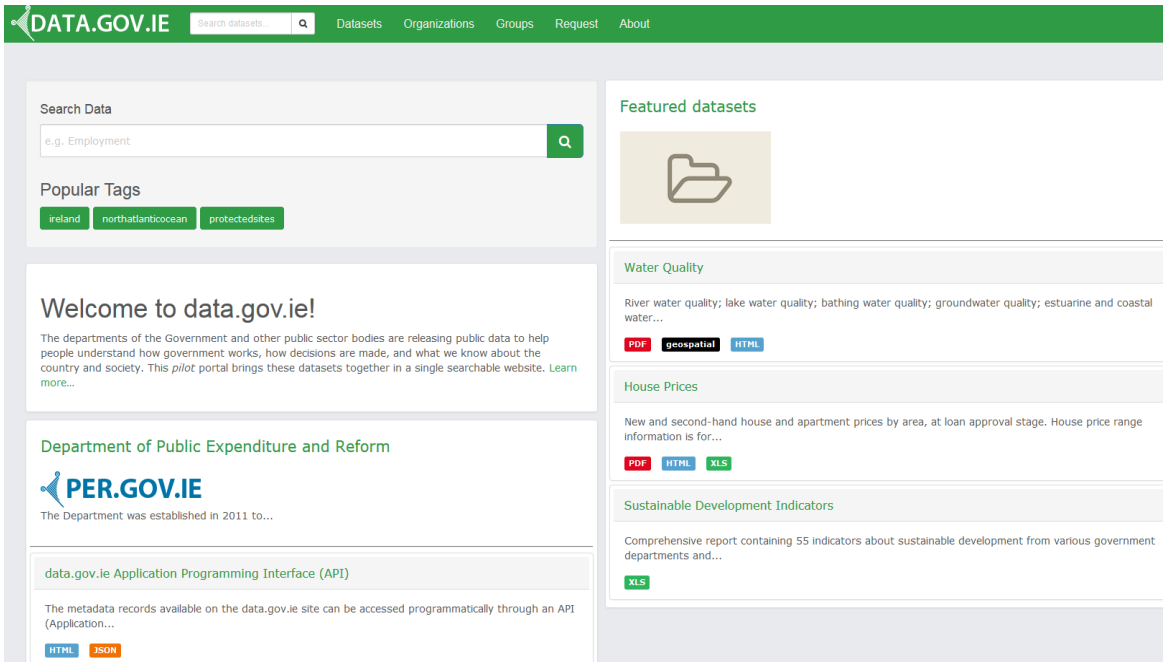

Figure 1. Department of Public Expenditure and Reform Open Data Portal

Open data in Ireland is being led and coordinated by the DPER Government Reform Unit and the Chief Information Officer. The new portal was created by Insight Galway using the CKAN platform and it marks an important milestone toward meeting the following government objectives to:

- enhance openness and transparency,
- strengthen public governance and
- yield economic benefits.

The following information products were delivered by Insight:

1. Best Practice Handbook
2. Data Audit Report
3. Roadmap
4. Evaluation Framework
5. Open Data Publication Handbook
6. Data.gov.ie

These information products were delivered as per the DPER Request for Tender (RfT) and according to the following high-level requirements:

- Provide advice and support to facilitate the implementation of key objectives and actions related to the open data initiative.
- Identify and align with national requirements best practice and international standards for open data;
- Help specify short- , medium- and long-term objectives for open data in Ireland;
- Identify potential opportunities for Ireland to design its open data drawing on the experience of what has worked and has not worked in other jurisdictions and also drawing on the scope for positively differentiating open data in Ireland;
- Establish an inventory of datasets currently available in the public domain as well as assist in identifying priority areas for the publication of datasets in the future; and
- Develop a roadmap for how Ireland could achieve its objectives in the area of open data.

In the spirit of collaboration and openness, after the July 22nd launch the DPER solicited input on the following questions from public sector bodies that have expertise and a stake in the initiative:

- What do you see as the priorities for open data and next steps?
- Are there any issues/concerns that need to be addressed in standardisation and publication of data?
- How we would maintain this type of group going forward given the level of expertise and experience in the room today – maybe meet as a forum 3 times a year?
- Indicate whether you are willing to participate in a smaller group and work in an engaged way to support DPER on particular aspects and on the development of an Open Data strategy in the coming months?
- How we should engage with civil society, business interests and other users to ensure we focus our attention on high value datasets that will be used so that potential economic, social and political benefits can be realised?

Finally, the DPER solicited input from the public, but unfortunately submissions were to be made to an email address in lieu of an open forum whereby others could see who and what was submitted.

The launch marks a turning point in Ireland's open data story. It affirms Ireland's international commitment to the Open Government Partnership (OGP) and it is also fulfils a precondition for Ireland to sign onto the G8 Open Data Charter whereby signatories commit to:

- Open Data by Default
- Quality and Quantity
- Useable by All
- Releasing Data for Improved Governance
- Releasing Data for Innovation

OGP membership, signing onto the G8 Charter and the launch of the Open Data portal, are positive steps toward the advancement of public sector reform and Minister Howlin, DPER and Insight are to be lauded for moving this plan forward.

The rest of this Working Paper examines the DPER Open Data plan, the research documents and their recommendations, and provides answers to a series of questions DPER officials posed after the July 22, 2014 public sector consultations. It is also The Programmable City project's submission to the DPER Open Data Consultation.

1. Good Governance or Open Data?

After examining open data reports, plans, strategies, indicators, and portals from around the world since their first official appearance under the Open Data banner in and around 2009, it is clear that open data initiatives should not be stand alone programs. Instead they should be a key component of government information management (IM), record management, IT and national spatial data infrastructures (NSDI), and be integrated into government IT infrastructure programs and shared services. In other words, open data should be part of a coordinated data and information dissemination strategy, which should also include publicly funded research data, scientific data, data from the humanities, and other qualitative data, in addition to administrative and government data.

Data, like government documents (GovDocs), are a way for government to communicate what it is doing and how it is doing it, and how it is doing, with its citizenry, its public, its public sector and the international community such as the OECD, World Bank, EU and other transnational, national and multi-national entities, be they private sector or civil society. And like GovDocs these need to be curated, managed, analysed, visualized, and disseminated on their respective government websites, discoverable in portals or by way of linked data, and be deposited in research and public libraries to ensure broad access. These also need to be preserved and archived as they are the official records upon which state decisions are made and programs managed. If we believe in both the ideals of democracy, democratic engagement, and the often contested concept of a knowledge based economy, then open data becomes a natural extension of those ideals.

It is unfortunate, however, that most Open Data strategies focus less on those ideals, and instead on the false promise of 'yielding economic benefits' and the 'release of data for innovation' in lieu of simply focusing on reaping the benefits of the good governance of a democracy. There may well be some economic return on the release of data, but those will most likely be in the way of savings and efficiencies resulting from time saved searching, cleaning, negotiating access and better management, rather than in the creation of 'killer apps'. Although, the release of core or framework data (e.g., postal codes, road networks), do provide the base information needed for many other sectors, and the financial returns of these are very significant (e.g., OSI, satellite imagery). In other words, the focus should be less on 'high value' datasets, and more on the production and maintenance of good quality public sector, administration, geospatial, research and scientific datasets, and conducted in the spirit of openness and collaboration with the public. For example, the publication of laws combined with the implementation of auditing practices, among other administrative reforms, were not put in place to promote innovation or economic returns, but doing so created the incentive to do business as rules are clear, the chances of corruption are reduced, and the public sector is held accountable, levelling out the playing field. These have been uneven and not always perfect reforms, but at least in western democratic administrations there is a nod toward openness and accountability. It remains a work in progress, but what is important is the incentive structures. If the focus is 'innovation' and 'high

value' then we may not see the release of datasets that promote a more egalitarian and fair minded society. Worse, we may get disappointed in the low economic return on open data initiatives and forgo them in the long-run because we measured the wrong outcomes.

The datasets about important societal issues such as homelessness, poverty, crime, education, environment, data to inform social and environmental impact assessments and planning, and so on might not fall into 'high value' categories, although their inclusion in indicators of well-being, quality of life and happiness would makes these of very high value. Sadly, open data portals do not contain the data required to construct most well-being and quality of life indicators, as those datasets cannot be used for commercial purposes. This will also mean a focus on framework or core data (i.e., postal code, road networks, topographic data, political boundaries and catchment areas, small area data) upon which data based narratives are built and made meaningful.

Open data is also no guarantee of a more evidence-informed, deliberative and engaged society unless there are broader legal reforms, political openness, transparency, education, capacity building, a critically numerate journalism, polity and society, and the welcoming of real and meaningful citizen engagement. Those are the ideals of open data, but most often and unfortunately the least focused on aspects of them. Even the principles in the Open Definition miss these. The unique moment for open data is not the technocratic sublime found in the California Ideology, in the dreams of open source libertarians, the niche marketing potential of big data companies, or the transparency banners of politicians, it is the possibility of governing together to improve our societies, and, that first and foremost should be the drive.

Recommendation:

- Open data should be a natural extension of good governance strategies and not stand alone programs.
- Open data should be a key component of government information management (IM), record management, IT and national spatial data infrastructures (NSDI).
- Open data should be part of a coordinated data and information dissemination strategy, which should also include publicly funded research data, scientific data, data from the humanities, and other qualitative data.
- Focus on the production and maintenance of good quality public sector, administration, geospatial, research and scientific datasets and less on commercialization, innovation and 'high value' datasets.
- Focus on data that have societal and environmental value, and also on core/framework datasets upon which other datasets can be integrated into.

2. Open Data, Records & Information Management and Thinking more Critically about Data

If the focus of open data is good governance, then open data becomes part of government information management/information technology (IM/IT) systems, and records management. If government digital records (i.e., data and related information products) are well managed, then opening them to the public is easy. The weakness of most international open data strategies, is their lack of integration into broader IM/IT systems and shared services infrastructure streams, which in the long-run, will makes them unsustainable and potentially turn them into one more institutional information 'stove pipe'.

The spirit of the open data movement internationally has been inspiring, and the work resulting from it impressive. It is however time for it to mature, and for those of us involved to think more systemically and recognize that science, geomatics, librarians and archivists, IM/IT have been at this for a while. In fact they have been doing all the things that open data enthusiasts aspire to, although not always in an open way.

It is also no accident that the first and normally the best and most abundant datasets found in open data portals globally are geospatial. Geography as a discipline, by its nature as the study of human interaction with the environment, is integrative. People and the environment are cross thematic disciplines. To tackle any issue in the economy, society, and environment requires collaboration with multiple communities of practices, experts, scientists, planners, and jurisdictions. This epistemic community has for more than a century been standardizing their terminology and practices, including precision time, longitude and latitude, navigation.

Geomaticians were early adopters of computerization, the first GIS having been created in 1964 as part of the Canada Land Inventory. Furthermore, environmental issues do not respect borders; we need only consider poor air quality, water ways, deforestation, transportation networks, disease, war, natural calamities, and etc. As a result there are indicators, maps, framework data, semantic ontologies, structured vocabularies, standards, metadata, catalogues, APIs, file sharing and conversion protocols, web services, shared hard infrastructure, interoperability standards, to name the most common practices, which ensure that at any given time data can be brought together to create a multilayered single views of spaces. National governments have been honing the skills of their public servants and scientists for quite some time. It is not perfect, but there are some stellar examples where it works, especially in the realm of spatial data infrastructures. The geospatial community in Ireland, for example, have been advocating for an National Spatial Data Infrastructure (NSDI) for some time, and have already build some inter-departmental and cross-sectoral collaborations, as seen in excellent portals such as Geoportal, AIRO, and the ISDE. In addition, the Ordnance Survey of Ireland (OSI) with the implementation of Prime 2 and the proposed formation of Tailte Éireann (consisting of a partnership between Ordnance Survey of Ireland, the Property Registration Authority and Valuation Office) are examples of this type of framework data, big data and integrative thinking.

The following is a list of elements that archives, records management plans, NSDIs and government IM/IT initiatives in general have in place to be successful and sustainable:

- A governance structure with a mission, vision, mandate, aims and objectives
- An organizational entity responsible for delivering the work
- Policy, legal, organizational, semantic and technological interoperability
- Skilled human resources
- Long term hard infrastructure management plans
- Technologies
- Multi-sectoral, -jurisdictional, -departmental, and -disciplinary collaboration and partnerships
- Standards
- Dissemination processes - access and discovery (i.e., portals, maps, catalogues, linked data, unique identifiers)
- Open architecture
- Open source
- Open licences

- Data quality
- Security
- Sustainable funding
- Evaluation

The data preservation community has developed a standardized checklist which includes the above and more, including sustainable technology, both software and hardware long-term chance plans.The following two primers illustrate how the geospatial community is implementing some of these practices.

The following principles from the Canadian Geospatial Data Infrastructure Vision, Mission and Roadmap - The Way Forward should resonate with the Open Data community.

1. ***Open:*** enables better decision making, the CGDI [Canadian NSDI] is based on open, barrier-free data sharing and standards that allow users to exchange data.
2. ***Accessible:*** allows users to access data and services seamlessly, despite any complexities of the underlying technology.
3. ***Evolving:*** the network of organizations participating in the CGDI will continue to address new requirements and business applications for information and service delivery to their respective users.
4. ***Timely:*** the CGDI is based on technologies and services that support timely or real-time access to information.
5. ***Sustainable:*** is sustained by the contributions of the participating organizations and broad user community and through the infrastructure's relevance to these groups.
6. ***Self-organizing:*** the CGDI enables various organizations to contribute geospatial information, services and applications, and guide the infrastructure's development.
7. ***User and community driven:*** emphasizes the nurturing of and service to a broad user community. These users, including Canadians in general, will drive the CGDI's development based on user requirements.
8. ***Closest to source:*** maximizes efficiency and quality by encouraging organizations closest to source to provide data and services. Thereby eliminating duplication and overlap.
9. ***Trustworthy:*** is continually enhanced to protect sensitive and proprietary data. The CGDI offers this protection through policies and mechanisms that enable data to be assessed for quality and trusted by users.

The components of the DPER/Insight adopted Open Definition focus more on data as objects and the technical qualities of a dataset and their attributes. Unfortunately, the Open Definition is a technocentric, computer science, and object-based approach which does not situate data in a broader IM/IT and records management context, thus masking the need for a broader systemic approach and the reliance on the soft and hard components of infrastructure. Instead, the open data community think of data in mechanical terms, as input into apps or big data analytical processes or as things to be found in a portal and not as government records that need to be managed, nor as national knowledge bases or memory systems.

It is important to remember, that when governments manage records, they do not do so for monetization, innovation or commercial reasons, they most often do so because there is a legal requirement to do so. Records are part of the government's memory and accountability system. When governments produce censuses, or map their natural resources, they do so because they

need to manage their nations, its resources and population, territory, economic and health matters. Data are therefore produced and used to steer the course of the nation, not for making apps. Data need to be thought of in a broader context, and a new discipline is emerging called critical data studies, with Kitchin's (2013) *Four Critiques of Open Data* exemplifying that mode. Thinking critically does not mean slowing or stopping work before all has been considered, but it does mean thinking more broadly as one is doing the work and continually checking in with stakeholders and whether or not what is being done is in the public interest.

Figure 2 from the Digital Curation Centre, promotes a life cycle management approach to the curation, preservation and publication of digital data assets. A life-cycle approach is common and governments have adopted it as part of government records management or information management policy for quite some time. Numerous institutions manage their data assets in this way (see guidelines here). The UK Data Archive provides another more simplified version. It is suggested here that adopting a life-cycle management approach alighn open data with the *Access, Use and Reuse* components of a record's life-cycle. TheOpen Definition fits well with these.This approach also aligns with conducting a data audit from a curatorial perspective, and with the practices a trusted digital repository standard and certification processes.

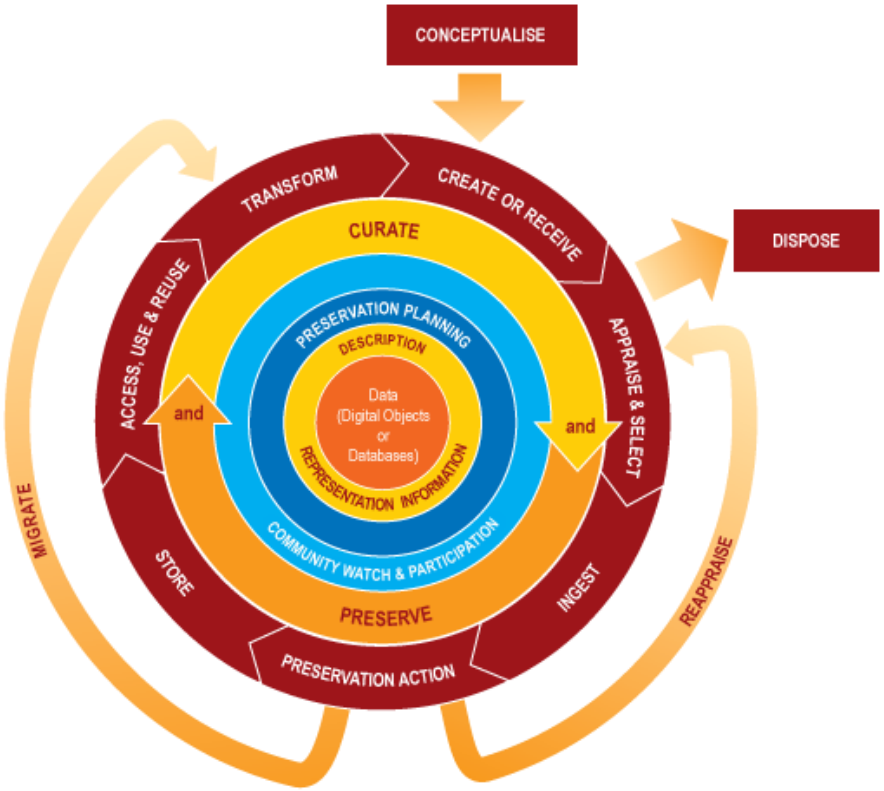

Figure 2. Life Cycle Management of Digital Datasets (Digital Curation Centre)

Another noteworthy example is the Canadian Federal Geospatial Platform, which includes open data, positions geospatial data as part of open government, incorporates the Canadian Geospatial Data Infrastructure and is part of a government renewal strategy (see Figure 3). This is a comprehensive socio-technological strategy that delivers digital data to citizens and manages them as assets. Like the trusted digital repository, infrastructure is a key component of this platform. Most open data strategies make no reference to the architecture, software and physical

infrastructures, cloud computing storage, high speed access and etc. which underpin the sharing, management and preservation of data.

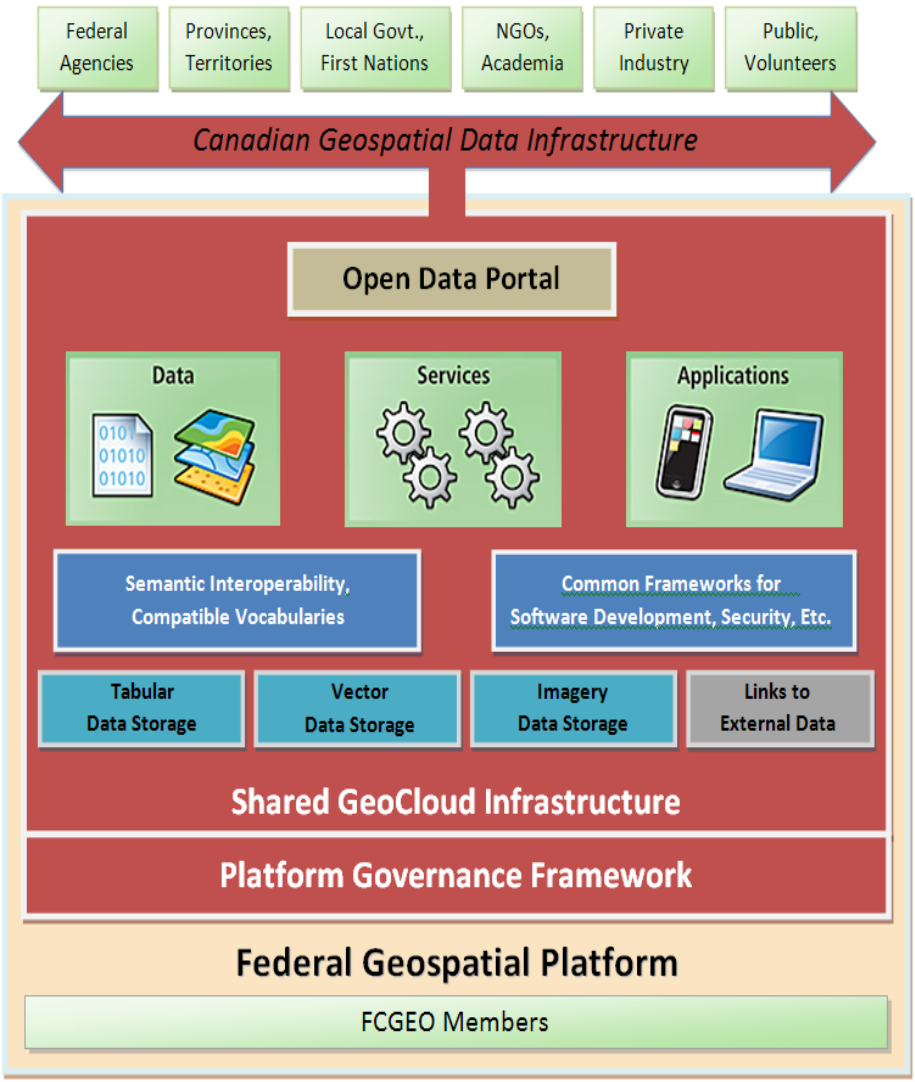

Figure 3. The Federal Geospatial Platform (Natural Resources Canada 2013)

If the government of Ireland considers an IM/IT, NSDI and life-cycle approach to managing its digital assets, then it will find that there are many very skilled people available in those often overlooked functions of government (i.e., IT managers, librarians, archivists, information studies, the geospatial community, scientific community), that could be mobilized into the plan.

The DPER/Insight Best Practice Handbook ecosystem diagram (p.9; see Figure 4) illustrates the common open data approach. It is based on object-oriented and programming thinking and frames data as technological objects with attributes separate from the organizations, institutions and knowledge systems within which these are situated. It is a helpful way to diagrammatically visualize how the report is structured, but this is not an ecosystem. It is analogous to thinking about the qualities of data and not data quality. The CGDI Principles listed earlier, exemplify ecosystem thinking, as does the DCC life-cycle management diagram as they take a more holistic approach to data and consider the inter-linkages and overlapping nature of data processes rather than their constituent parts. Alternatively, the Programmable City project thinks of data assemblages, which is a systemic approach to examining data from a critical data studies

perspective. The Federal Geospatial Platform diagram more so resembles an architectural framework, and can be conceptualized as a bioregion within a larger information ecosystem.

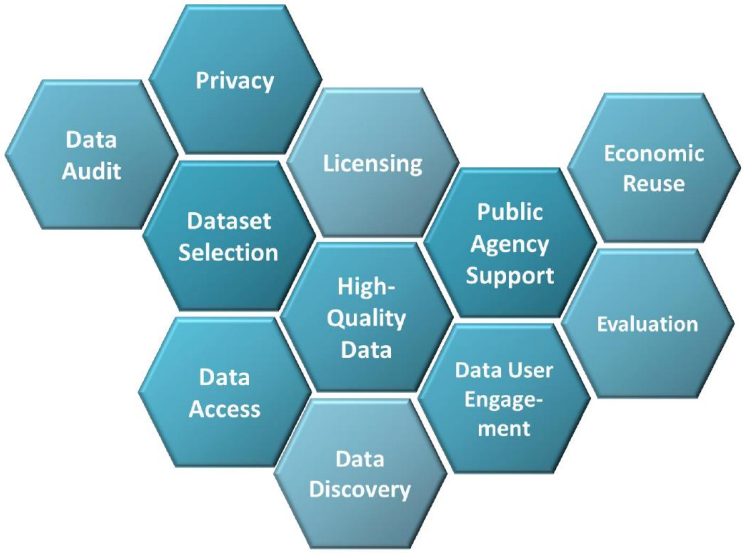

Figure 4. Open Data Ecosystem (DPER / Insight Best Practice Handbook 2014:9).

Recommendation:

- Consider Open Data as a good governance strategy and as part of data and information management
- Integrate Open Data into IM/IT, Shared Services and integrate with the NSDI
- Data infrastructures are critically important, Open Data should be considered with cloud computing, high speed internet, and hardware and software.
- Open datasets should be thought of as government records (data & information) and should be managed accordingly
- Adopt a life-cycle and data curation approach to the management, preservation and dissemination of Open Data datasets
- Implement the NSDI and consider the CGDI principles for the NSDI and for Open Data in Ireland
- Critically reflect on data more broadly and not just as objects at the end of an information pipeline.
- Consider evaluating the contents of an open data portal to see if these can be used to construct indicators of well-being and quality of life.

3. The DPER / Insight Roadmap and the Best Practices Handbook

The proposed Open Data Roadmap provides short, medium and long-term objectives. It also assembles, into one document, the recommendations from the Best Practice Handbook. The following examines some of the items in those documents.

3.1. Governance

The following, combined with the above is a response to Section 3 Governance of the Roadmap. It is suggested here that for open data and open government programs to succeed, there is a need for a strong public interest mandate, vision and mission, and clear objectives against which performance can be evaluated. In order to implement these a dedicated institution, such as a secretariat, division, or a special unit in the public sector, is required. Once open data becomes a normalized part of information management in the public sector, most likely in a 5-10 year timeline, that office may no longer be needed. That office would need to be strategically situated in the administration where it would have coordinative, integrative and some directive powers to collaborate with all departments at all levels of government in order for it to effectively carry out its work. We support the Roadmap's recommendation for the creation of a new Open Data Officer, and in addition, each department should also have a public servant occupying this type of role. Their job would be to coordinate data audits, metadata, etc. within their respective departments and respond to queries and integrate practices within IM/IT.

Success will also be contingent on a dedicated budget and a team of personnel skilled in international, national and local public policy and law, collaborative governance and public engagement, multi-sector partnerships, and technologically proficient in data management, programmers, administrative, public, spatial and scientific data, including experience in technologically mediated engagement, venture capital data analytics, and so on. There are people throughout the public sector (government, academia & cultural institutions), at all levels in Ireland with the will and skill to be a part of an open data/open government entity, many of which could be recruited and seconded from successful and existing data dissemination programs for a pre-determined amount of time. The small list of initiatives in a Programmable City Open Data is a good place to start looking, the proponents of the NSDI is another, as is current records management HR listings within the public sector, and of course the list of participants at both the DPER's public sector and civil society consultations in July and September of 2014.

If such an open data/open government entity were to be created, it would be responsible for convening the DPER proposed Open Data Board (ODP) and the Steering and Implementation Group (SIG). It would also be able to convene thematic working groups (WGs) of known experts to develop plans, policies, standards, technologies, processes, applications, audits, (i.e. open data components in the Best Practice Handbook), citizen outreach and objectives as outlined in the Roadmap. The WGs would feed into the SIG. WGs would be temporary, and experts would develop strategies and implement them, and align them with records management and IM/IT processes, as well as with Shared Services and the NSDI. Once those processes are in operation the WG would dissolve. The open data/open government entity would be responsible for helping actors in the SIG & WGs operationalize their work. Geoportal, ISDE and AIRO are currently doing this type of work while the Digital Repository of Ireland has developed an organizational structure to carry out their cultural data preservation program. Concurrently the DRI has produced some excellent resources directly related to open data. Technical, policy, legal, thematic, curatorial, and managerial expertise, from these groups should be sought, as should expertise from open data veterans (e.g. Fingal Open Data). A rough schematic of this structure would look something like Figure 5:

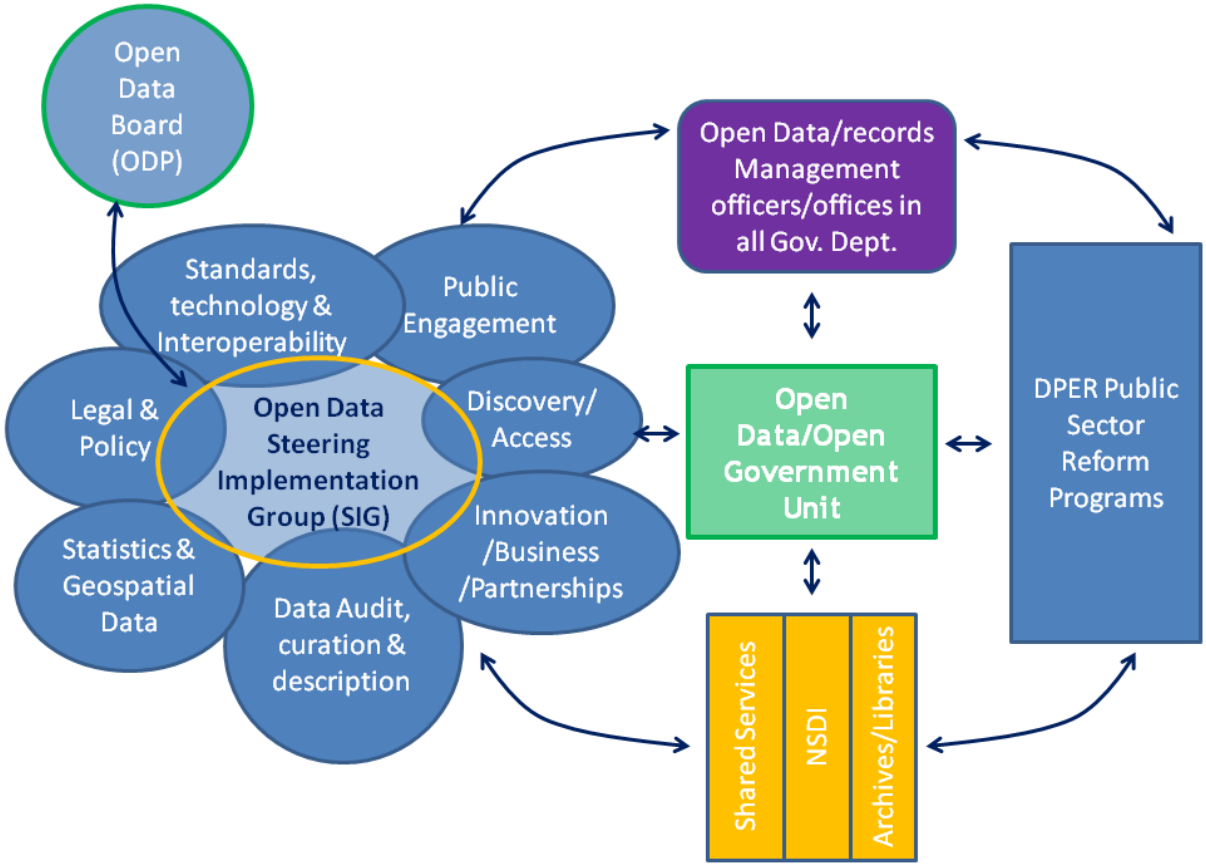

Figure 5. Schematic of an Open Data Organizational Structure for the Government of Ireland

The proposed DPER ODB and SIG are good, and details about their composition and mandate would require further discussion, however, it cannot be overstated that it is doubtful that a successful open data/open government plan can implement what is in the Roadmap without a dedicated public sector entity with a mission, mandate, vision, powers, budget and dedicated human resources. In addition, for the program to be sustainable, it should be integrated with the national IM/IT, Shared Services, and NSDI, and eventually become a natural aspect of those government functions. The SIG on its own, it is believed, would not have the capacity to do this, but could steer WGs in collaboration with the open data/open government entity, and it would receive high level strategic guidance from the ODP which would have the responsibility of ensuring that this entity is politically and financially supported.

Chapter 13: *Best Practice Standards for Supporting Public Bodies Guide* in the DPER/Insight Handbook provides some useful recommendations regarding leadership, policy, financial and capacity building and these are reiterated in Section 4.12, *Supporting Public Bodies,* of the Roadmap with a focus on the central government and each public body. Those recommendations combined with the recommendations and the organizational structure just discussed could lead to a more inclusive, informed open and sustainable open data strategy that becomes a natural extension of government services. Existing skills from the public sector could be mobilized as no extra expense and temporarily dedicated to this activity.

Finally, the ODB and SIG terms of reference, (see Roadmap, Section 3) should be revisited to reflect mission, vision and mandate of the Open Data Strategy. Members should be appointed

by peers, it should be an open process and selection should be based on skills and explicit functions. Ministerial approval is helpful, but these should not be solely political appointments.

Recommendations:

- Develop an open data public interest mandate, vision and mission, and clear objectives against which performance can be evaluated.
- Reconsider the organizational structure as per the schematic in Figure 5.
- Reconsider appointments on the SIG to be expertise and skills based and less political, and that appointments be made by peers.
- Create an open data institutional entity that will operationalize the work of the ODB, SIG and Working groups and integrate these with other government programs.
- Open data officers should be appointed in all government offices.
- Create temporary expert working groups to develop and implement infrastructure wide practices.

3.2. Legal, Policy and Ethical Framework

Public administrations are mandated by law to manage their records, and normally this is done under the guise of legal, policy and ethical frameworks. In some cases like a census, archives and land surveys, there are explicit pieces of legislation that govern how those are to be carried out. There is also overarching legislation that governs all government data collection and dissemination. Data protection, PSI Reuse License, and Freedom of Information are examples of overarching laws. There are also normative and ethical practices which frame how research data are to be collected and archived (see Irish Qualitative Data Archive). A data and information legal and policy framework would take all of these into consideration.

The management and dissemination of open data would have to adhere to the same legal, policy and ethical framework as would all data produced by the public sector. The DPER/Insight reports only address these in a piecemeal fashion. For example, data privacy and licensing are mentioned in the Roadmap (Sections 4.4 and 4.5) and in the Handbook (Chapter 8 and 9) as was Freedom of Information which was discussed in Sub-Section 3.3. As for directives and regulation, the EU PSI and its Irish instantiation are discussed (See sub-sections 3.5 and 3.5.1 in the Handbook), as is the INSPIRE directive in the context of metadata practices, data audits and the National Spatial Data Infrastructure. Finally, the Aarhus Convention and the Access to Information on the Environment regulation was also discussed within the Geospatial Community section. Ethics are not discussed at all, nor are aspects of contract law. These and more should be examined in more detail, and combined with specific instructions in the Open Data Publication Handbook. Law, regulation, ethics and policy are not attributes of a dataset, they are the conditions which govern the conduct and how government manages its data and information assets. This approach is part of the life-cycle management of data and the NSDI.

Along with references to policies on pages 40-42 of the Handbook (Government Policy and the Geospatial Community), the following instruments should be considered: procurement processes, memorandums of understanding, licenses, records management policies, data retention policies, contracts, data sharing agreements, cost sharing, infrastructure sharing, work secondments, along with any other instruments that can be used to foster more openness, sharing and collaboration between and among institutions, how data are to be acquired and managed.

The following is a partial list of statutes that govern the management and dissemination of government data and records in Ireland as well as an Open Data plan:

- National Archives Act
- Statistics Act
- Intellectual Property
- Intellectual Property (Miscellaneous Provisions) Act
- Copyright Act (all its amendments)
- Trademark Act
- Patent Act
- Data Protection (all its amendments)
- Digital Hub Development Agency
- National Standards Authority
- Documentary Evidence Act
- Freedom of Information Act
- Public Records Fees
- Public Records Act
- Communications (Retention of Data) Act
- European Communities (Access to Information on the Environment), Regulation
- European Communities (Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE))
- European Communities (Re-Use of Public Sector Information)
- Health Information and Quality Authority
- Chief Scientific Advisor, Scientific Policy Documents of Ireland
- And any others which govern government records

The *Analysis of Intellectual Property Activity in Ireland Based on Existing Data* report resulting from the RfT in the spring of 2014 should also provide some insight as to the state of law and regulation in Ireland. There are a number of academic and civil society organizations in Ireland focusing on aspects of these issues: Creative Commons Ireland, the Copyright Association of Ireland, the TheStory.ie, and 1709 Blog, etc. Also see the Digital Repository of Ireland's resources and fact sheets. Finally, tools, processes and guidelines to anonymize and aggregate data should be designed to enable public administrators to disseminated data in accordance with laws, policies, directives and research ethics.

Recommendations:

- Develop a data and information legal and policy framework with open data as a component of it
- Conduct an inventory of collaborative and data sharing instruments (e.g., MOU, procurement contracts, data sharing agreements, etc.)
- Assess the outputs of the *Intellectual Property Activity in Ireland Based on Existing Data* report resulting from the RfT in the spring of 2014
- Conduct an inventory of all laws, regulation, policies and directives that would govern how data are collected and disseminated
- Develop a set of explicit legal, policy and ethical guidelines for the management of public sector data and open data based on laws, regulation, directives, policies and practices in Ireland for public sector officials

- Include these guidelines as part of the data dissemination decision-making tree that will be discussed later in this document.

3.3. Public Engagement

Government reform, in the way envisaged by DPER, should also include meaningful public engagement. Public engagement is a key component of open data and open government yet engagement is absent from the high level components of the DPER open data plan and it was not a specified requirement in the RfT.

Open data is part of good governance, and is associated with transparency, accountability, innovation, and the promotion of evidence based decision making within the public sector. Open data is also about deliberative democracy, which means receiving input from, being responsive to, and engaging with actors outside of government, such as the public, civil society organizations, the media, academia, and the private sector. Releasing data is in many respects the easy part, transforming the public sector, whether in Ireland or abroad, to meaningfully embrace informed and deliberative democracy and co-governing with citizens is much more difficult. This is cultural change. DPERs public sector reform programs are creating the mechanisms to institute broad based cultural change on many fronts, and it would be even better if engagement were more explicitly stated here.

More work is required to better understand how to technologically mediate public engagement, to conduct outreach, carry out consultations, to have face to face meetings with the public, to build capacity with the public sector to do so, and also to build the capacity of citizens to engage in the public interest rather than self interest is important. Organizations such as Engagement Lab, Change Camp, tools for public consultation such as those created by Open North, civic innovation institutions such as New Urban Mechanics, or the work of Code for Ireland and thematic hackathons, such as Charity Hack Day and Open Government Jam by the Open Knowledge Foundation Ireland, or BC Apps4ClimateChange or Hacking Health are examples to draw inspiration and expertise from. None of these are perfect on their own, but could be part of a multi-pronged approach with methods aimed at achieving a particular target on a particular issue. For example, AIRO has been conducting Data for Decisions workshops with public officials which help build capacity in the public sector to better understand how to use the data officials are responsible for, clean and map them. There is also expertise in the academic community with regard to user needs analysis, urban planning design charetes, surveys, and so on. While those involved with environmental and social impact assessments have developed consultative practices that can be considered here as an engagement plan.

An Open Data strategy can have an information push and an information pull approach. Crowdsourcing and volunteered geographic information (VGI) are examples of information pull approaches, which can be also considered as another form of public engagement. The Environmental Protection Agency in Ireland has been doing this form of work in a citizen science framework, while the Government of Victoria in Australia updates its maps this way with a Notification Edit Service, while the government of Canada has produced a VGI Primer. Increasingly, governments are considering these types of data as a way to reduce data collection costs but also as a way to engage the public. The use of Open Street Map for overseas development and disaster relief are classic examples of this.

Engagement needs to be meaningful and lead to action and change. More importantly, the impetus should be toward simultaneously improving the health and well being of Ireland's people, its environment and its economy, and this means acting in the public interest. Open data

and open government can be a test best for that kind of meaningful engagement, one that is evidence based and beneficial to Ireland as a whole and not just specific interest groups. However, open data on its own is not enough.

Chapter 14: *Best Practice Standards for Engaging Data Users* in the DPER/Insight Handbook provides a good overview of issues pertaining to engaging open data users with regard to its open data portal and the data it contains. Opening government more broadly, is another matter that means enabling government to address issues derived from the analyses resulting from open data, thus evidence based decision making and a more informed type of deliberative democracy.

Chapter 14 of the Handbook discusses potential data users; we suggest discussing the interests and requirements of current data clients while also targeting potential users. The public sector already has data clients and there are a number of partnerships, data sharing agreements - formal and not - in place, as well as procurement agreements that are not necessarily in the public interest and so on. Those arrangements need to be examined, and clients need to be brought into the engagement process.

Chapter 14 also focuses heavily on demand-driven data, and the report's overall focus is on the G8 high value datasets (see section 6.2 in the Technical Annex). We agree that this is important, however we caution the DPER to consider that what an app developer may want, or that a focus solely on data that can be commercialized, are not comprehensive approaches. The government produces data to govern, and it is these data that can be put to greater use. Civil society groups wishing to work with data on homelessness and social housing, or a village wanting to know about the quality of its ground water, or a business association wanting to know how many bankruptcies have occurred in its jurisdiction are valid reasons for the release of datasets, irrespective of how popular those datasets may be. Often it is these that can positively improve people's lives and their livelihood directly. Furthermore, many geospatial datasets would have a finite number of users; however, those few users are in high impact decision making areas such as transportation, agriculture, planning, and public health. Weighing high demand versus impact of a few is a difficult balancing act. If the mandate is open by default, then those not so popular datasets, including those that may cause controversy, also need to be shared. In addition, data and information requested from the FOI processes should also be considered as demand-driven data and those should be released.

Recommendations:

Along with creating an Engagement WG, the recommendations stated above, and extending the recommendations in Section 4.13 of the Roadmap, the following should be considered:

- engage with stakeholders on developing the mission, vision and mandate for the Open Data strategy.
- Engage with stakeholders to shape how an Open Data roadmap and strategy could look.
- Engage with, study, build upon and harmonize the Open Data strategy with existing public sector data dissemination programs.
- Review and assess existing technologically mediated engagement tools and social media applications in other jurisdictions.
- Public sector officials and departments should develop processes and be receptive to evidence based public input into public policy and planning, and learn to solicit feedback from the public in a useful and educated way.

- Consider crowdsourcing, VGI and citizen science as a public engagement strategy

4. Data Curation or a Data Audit?

The DPER RfT requirements were very narrow in scope and did not include conducting a detailed audit of a public sector entity to assess what type of datasets it produces and how these could be published. The conduct of the data audit was guided by the following:

- Specify criteria and guidelines for building a credible data audit toolkit that could be used by public bodies to enable them to carry out an audit of their datasets for potential publication;
- Identify existing published datasets for consideration for inclusion on the portal and
- Provide an assessment of sectors where new datasets should be audited in order of priority for potential inclusion on the portal on the basis that their publication in open data formats and have the potential for creating a significant economic and/or societal impact.

In the time DPER allotted to carry out this task, these requirements were reasonable, although to some extent, as stated in the Handbook, the community already knew what data were already published and these have already been made available in other existing portals. An audit of what is not published would have been much more insightful. Furthermore, many datasets have been published but are not in open formats, nor disseminated under and open license, and these too would not have been captured in the machine audit conducted. The DPER/Insight Data Audit Report, Chapter 5: *Best Practices for Carrying out a Data Audit* of the Best Practice Handbook and section 4.1 Carrying Out a Data Audit in the Roadmap discuss the details of the data audit and its limitations.

The Insight Data Audit was carried out as follows:

1. Web crawl of government websites and state bodies selected from the PSI website with the exclusion of local authorities. The web crawling tool was directed to search a selection of technical key words and file formats.
2. Reviewed the data in two catalogues StatCentral and the Irish Spatial Data Exchange (ISDE).
3. Manual review of datasets resulting from the webcrawl that were considered of high value according to the G8 Charter (see section 6.2 of the technical annex of the Open Data Charter).

This was a machine-driven technocentric approach to a data audit which missed many datasets, some of the limitations of which were discussed in the report. The results were narrower than what was stipulated in the RfT and as a result only 400+ datasets were identified. There was very little explanation for the adoption of the methodology applied, nor for the exclusion of local authorities, the omission of Geoportal, AIRO, Fingal Open Data, Dublinked, and the list of Irish Data Catalogues in Table 1. of the Handbook (p.34-36) for that matter, or for narrowly casting to only those listed on the psi.gov.ie website. There are also many datasets the government of Ireland publishes in international portals such as EuroStat which are of high value as these are

used to inform EU expenditures and policies regarding health, agriculture, the environment, energy that affect Ireland.

Furthermore, 'high value' was deemed to be datasets listed as priorities in the G8 Open Data Charter and other open data indicator projects, even though, the data in the portals just listed are considered to be of high business and administrative value to those who publish them, otherwise they would not be collected, described, managed, visualized, standardized and used to inform planning and governing decisions. Chapter 6: *Best Practice Standards for Dataset Selection*, the Handbook was broader and highlighted how the US Office of Management and Budget and the Government of New South Wales define high value, but it also overly emphasized the publishing of demand-driven datasets, which as discussed earlier overshadows the less demand but high impact datasets. The Handbook also reviewed the Open Data Census and Barometer, Danish Basic-Data Register and mapped these with G8 Charter and developed a list of Common High-Value Datasets (p.52-53). That mapping was a good piece of work. An examination of what datasets are most sold, requested or used at the moment, would also have been another way to assess high value datasets.

The data audit as it was set out was well executed and yielded 400+ datasets whose metadata were integrated into the data.gov.ie portal. It was however overly exclusive and an audit of assets would have been more useful, irrespective of how they were published as tables in webpages or embedded in databases or other types of systems, formats and licences. The G8 Charter and the other indicators systems are lauded for their objectives and advancing an open data agenda, but it is critically important to keep in mind that there is an epistemic disconnect between what open data enthusiasts consider to be of 'high value' and therefore 'monitor' in their indices, and what geomaticians, scientists, epidemiologists, economists, agronomist, meteorologist, administrators consider important for managing people, natural resources, industry and to govern. The Insight 'high value' list of data represents the aspirations of the former but not the reality of how things are done by the latter. The Handbook refers to a study conducted by the Sunlight Foundation Open Government Benchmark Study which reaffirms that view. The Insight data audit method was an expedient way of finding data that could easily be inserted into a portal, but it is not comprehensive. The data audit did however demonstrate what can be done in a short amount of time, and now the hard comes with maintaining the portal and adding data into it.

Having stated that, conducting an assessment of whether or not the public sector publishes the 'high value' datasets listed here is not a bad thing; however, suggesting that these are the only datasets considered to be of high value is problematic. Furthermore, if these data were discovered exclusively from the machine directed method discussed above, and not by conducting a manual online search, digging through reports, calling agencies directly, examining Eurostat etc., then many data would have been reported as missing. Also, some thematic expertise is required. For example, it was understood by the researchers involved in the production of Irish and the Canadian OKF Index, that it was critical to rely on the expertise of thematic experts, as someone with grounded knowledge in budget, spending or procurement may not be versed in air quality, water quality and transportation who know their domains and

can readily identify what is relevant and useful. People with domain knowledge need to be a part of the process of determining what is of 'high value' in their domains. Machines cannot do that.

The Handbook pointed to an excellent Data Audit Framework, very much in keeping with an information management (IM) approach to the life-cycle management of records and it is curious why this method was not applied here, for at least one department. The DAF has been scientifically developed, tested and used by the archives, library and scientific community. Many organizations in the EU have also used the DAF to assess their data assets and comply with the INSPIRE directive. It is a very good step by step guide to planning an audit, identifying and classifying assets, assessing how to manage identified assets as well as reporting and recommendations. The DAF method would also have to be aligned with a chosen geospatial and open data metadata description standard and in accordance with linked data practices.

The proposed Data Audit Framework will alleviate these short comings and it is highly recommended that the audit be broad-based and not based solely on certain key words, licenses and formats. It is also suggested that a curatorial approach be taken, as per the Data Audit Framework and the Life-Cycle Management diagram discussed earlier (see Figure 2). Curating means conducting a thorough data inventory or audit, and then carefully considering what has been discovered, the quality of the data, if they are described, and making informed decisions as to when and how to share these data. Curation comes from the cultural community, museums, archives, humanities and libraries, and they have developed excellent processes, as seen in the Digital Curation Centre, and there is merit in thinking of data as the nation's digital assets and as artifacts that represent culture.

The metadata Tables in the Data Audit Report (Sections 3.1-3.14), should also include the following, as these will speak to and attest to fit for use decisions and data quality:

- geocoded – yes or no
- geographic scale (i.e., small area, election districts, health regions)
- time scale (between when and when including the frequency)
- formats (irrespective of open or not, part of systems or geospatial formats that are open but unknown to the auditors)
- methodological guides and data dictionaries provided – yes or no
- data limitations (e.g., crime data are geocoded to the Garda station and not where the crime occurred)
- disclaimers
- data quality

The following is a list of datasets that were missed and are part of the G8 Charter datasets: Health Atlas Ireland, Elections Ireland (as imperfectly formatted as they are), Digital Repository of Ireland, The Irish Qualitative Data Archive, Eurostat, Irish Social Science Data Archive (ISSDA), Database Portal, ESPON Database, , CSO Commuting in Ireland, Commissioner of Irish Lights Data, Science Foundation Ireland data. Ireland produces many high quality historical and research datasets, as well numerous scientific, environmental and government datasets. An inventory of portals from all sectors, where the data are open or not is important,

as is a way to integrate their metadata to ensure ease with discoverability. Arguably, portals are an interim step; ideally these should be matched with linked data practices, at least for the discovery metadata.

Finally, the full results of any departmental data audit should be made public to enable the public to see what is available, and to evaluate openness against it. That list becomes a kind of shopping list for the public.

Recommendation:

- Adopt a digital data curation and life-cycle approach to the management of data and conduct the data audit accordingly
- Adopt the Data Audit Framework
- Ensure that the additional elements are added into the data audit (e.g. geocoded, scale, time)
- The high value approach to the selection of data should be reconsidered, and an evaluation of what current data 'clients' value, should be considered
- Recognize the limitations of a machine only audit, and broaden search criteria to include all data not just those in open formats and under an open licence
- Conduct a full inventory of portals and catalogues from all sectors in Ireland and integrate their metadata to ensure cross disciplinary discoverability.
- Publish the results of the data audit.

5. Data Dissemination and Publication

The DPER RfT specified the need for a document that would set out the practical steps to be taken by public bodies to facilitate the publication of data in open data formats, (e.g. the creation, formatting, integration, storage and publication of data) and Insight produced the Open Data Publication Handbook.

The Open Data Publication Handbook recommends doing an audit, but did not point to the excellent Data Audit Framework method discussed earlier. It also recommended: an examination of existing catalogues; that a machine led online audit be carried out; the conduct of an internal audit; the publishing of data according to the open definition and under a national Open Data licence which is currently not available in Ireland therefore under a PSI; to publish the data listed in the top international indices and to ensure that data protection laws are adhered to. It did not really explain in detail how to do these things.

Furthermore, the Open Data Publication Handbook stated that data should be published according to the 5 Star rating but did not provide any tools or guidelines on how to do so. There was also the recommendation that standardized metadata schemas be adopted but provides none, and did not discuss crosswalking existing portals with each other, with Dublin Core, or to the established ISO 19115 standard of the geospatial community. Geospatial metadata should not be downgraded to the open data standards, but open data metadata should at least adhere to the high level metadata of ISO 19115. Interoperable tools exist, and GeoPortal and the Irish Spatial Data Exchange (ISDE) are demonstrations of that. Another example is the

GeoConnections Discovery Portal developed by Natural Resources Canada. It is part of the Canadian Geospatial Data Infrastructure (CGDI) and is an open architecture, open source, interoperable portal, that adheres to standard metadata, open data, and provides multiple ways to access the data (i.e., text search, coordinate search, map search). The CGDI also produces a number of very helpful operational policies and standards resources that would be very helpful to DPER. The geospatial community is very skilled in Ireland and the Irish Organisation for Geographic Information (IRLOGI) should be consulted as should those involved with the NSDI.

The Publication Handbook also suggested that data should be published in open formats but did not specify what those were. Irrespective, as a first step, public sector bodies can publish data in the formats within which they are created and provide file conversion instructions or services, until which time legacy systems change. GeoGratis, the first Open Data portal in Canada launched in early 2000 did just that, and as systems were upgraded, new open formats, or de facto standard formats were used. Also, national archives often produce file format registries (e.g., UDFR, Library and Archives Canada) as they recognize as preservation institutions, that they may have to ingest data in multiple formats for their long term archiving and maintenance. An examination of these registries could be of use to DPER. The Digital Repository of Ireland also created a number of useful publications on this and related topics, and has a list of file formats used in the social science in its *Digital Archiving in Ireland: National Survey of the Humanities and Social Sciences*. The Research Data Alliance (RDA) is also tackling many of these issues and some public sector officials are members of the RDA as well as a number of other standards bodies such as the ISO, as would be the National Standards Association of Ireland (NSAI). Different domains publish data in different formats for a variety of reasons, and it is absolutely reasonable that administrative data be transformed into open formats as linked data and described with standard metadata. It would be more reasonable however, if some data, for quality, economic and labour reasons be made available as they are, and that users convert them on their own, until which time formats and new systems are in place.

The Publication Handbook also recommends the adoption of unique identifiers, but did not recommend how to go about doing so, and also that data should be made accessible mostly via APIs, but did not discuss that this requires an infrastructural shift for government and a 24/7 service agreement. While this is a useful approach, discussions on the role of government and APIs and how these are integrated with shared services should be had.

The following were not discussed in the Open Data Publication Handbook and often not discussed by the Open Data community:

- Interoperability – of data, technologies, portals, policies, laws, integration of existing portals (see GeoConnections resources)
- GeoCoding and geoenabling administrative datasets
- Data aggregation and anonymization methods (consider experts at the National Centre for GeoComputation at NUIM)
- Methods to create linked data (Insight NUIG)

- Federating metadata and pointing to collections of data in lieu of replicating data (Open Data strategy and records management approach)
- Distinguish publishing processes based on administrative data used to inform programs, government data such as census, and scientific/geospatial data and develop guides accordingly
- Ethical guidelines and policies regarding the dissemination of sensitive and research data
- Legal, policy and ethical framework with explicit guidelines
- Develop a decision making tree for public sector officials when it comes to determining what and how to publish data
- Develop peer learning environments
- Data quality and cleaning guidelines
- Data preservation
- General data management principals and strategies

Recommendations:

- It is highly recommended that DPER consider adopting the well established data curation life cycle management approach similar to the one developed by the Digital Curation Centre, and consider taking a data curatorial approach in lieu of a data audit.
- Adopt the Data Audit Framework for data curation as well as those developed by the Digital Curation Centre and consider developing an Information Management Directive which incorporates the ideals of Open Data, preservation and archives.
- Open data is not standalone, it is part of the data and information management of government records in general. Unfortunately, most open data initiatives are created as standalone processes separate from records and information management. DPER, could develop a data management and dissemination WG, and invite experts from the Digital Repository of Ireland, library and archives and information studies, geospatial community to help develop a comprehensive access, dissemination, data management and preservation plan for Ireland. That would be in keeping with the Roadmap which states that: "To maximize the benefits of open data for society synergies and collaboration between the different areas need to be established – between society, government, research, libraries and archives, but ultimately also establishing interoperability with private and business data sources".
- Create a decision making tree to help public officials determine what can and cannot be published. Figure 6 is an example to guide decisions on the management and dissemination of sensitive data (P.27 *Best Practices for Sharing Sensitive Environmental Geospatial Data*):

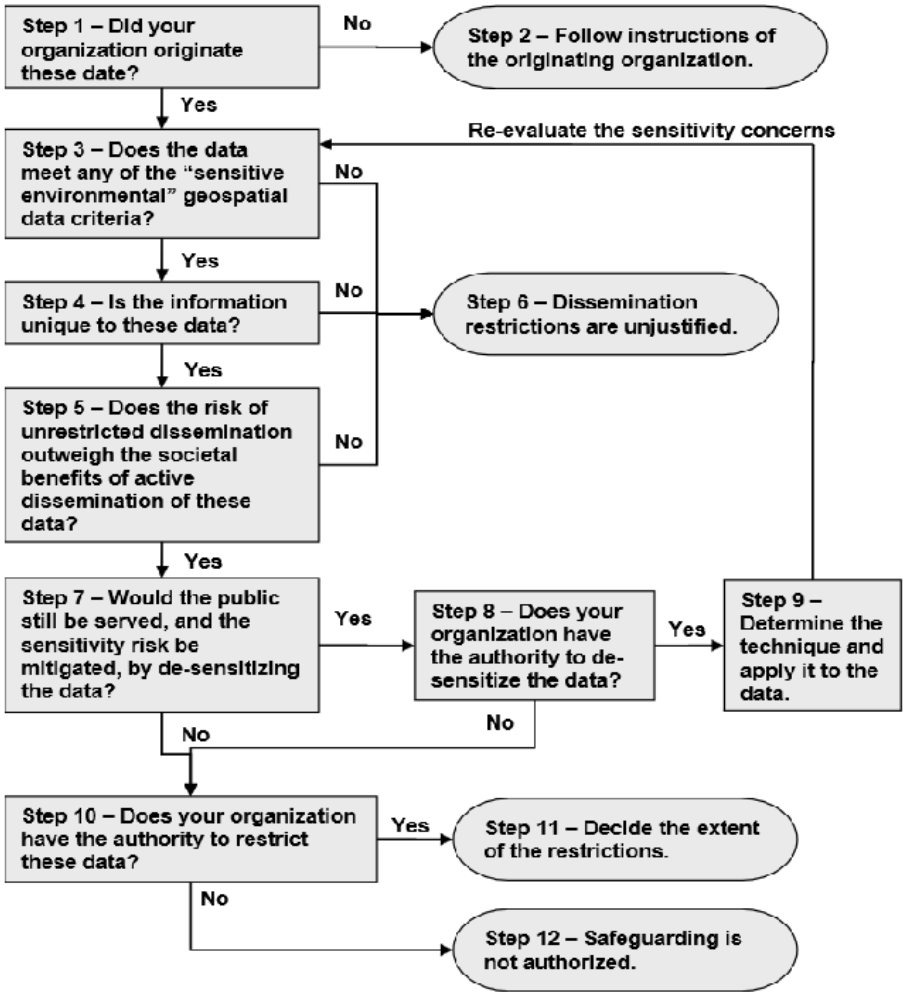

Figure 6. Decision Making Tree on the Publishing of Sensitive Data (Best Practices for Sharing Sensitive Environmental Geospatial Data Guide, Natural Resources Canada, p.27)

- The outcomes of the decision derived from the application of the open data publication decision making tree would then form the basis for the decision supporting not publishing a specific dataset by default.

6. Evaluation

An open data/open government evaluation process requires: mission and vision statements upon which a mandate can be developed to inform a strategy, along with a business plan, a suitable organizational structure and an entity that can operationalize that strategy. For it to be valid, the mission, vision and mandate should be developed in a well structured and facilitated public engagement process that would also include reaching out to current data clients and the main organizations that currently successfully disseminate data in Ireland. The outcome of that process in tandem with the submitted roadmap could then be used to develop WGs to advance the strategy in collaboration with the SIG. This could be a function of the Open Data Board (ODB), while steering the adoption of a particular evaluation process and its conduct would be overseen by the SIG in collaboration with the newly created open data/open government entity.

The DPER/Insight Evaluation Framework refers to Chapter 16: *Best Practice Standards for Evaluation*; the *Current State of Irish* Practice sections in Chapters 5-16 of the Handbook and the short, medium and long term objectives as well as Section 4.15 *Evaluating Impact* in the Roadmap.

The Open Data Barometer indicator project and the Open Data in Developing Countries (ODDC) Conceptual Framework are the main ingredients of the proposed approach in the DPER/Insight Evaluation Framework. The ODDC Conceptual Framework recommends a multi-leveled and multidisciplinary approach which includes a qualitative assessment of open data projects as well as an examination of research emerging from public policy, organizational theory, political science, economics and etc. The ODDC approach remains a work in progress and was developed as a research framework to study open data initiatives in developing countries. The components and the conceptual framework guiding ODDC research is illustrated in Figure 7.

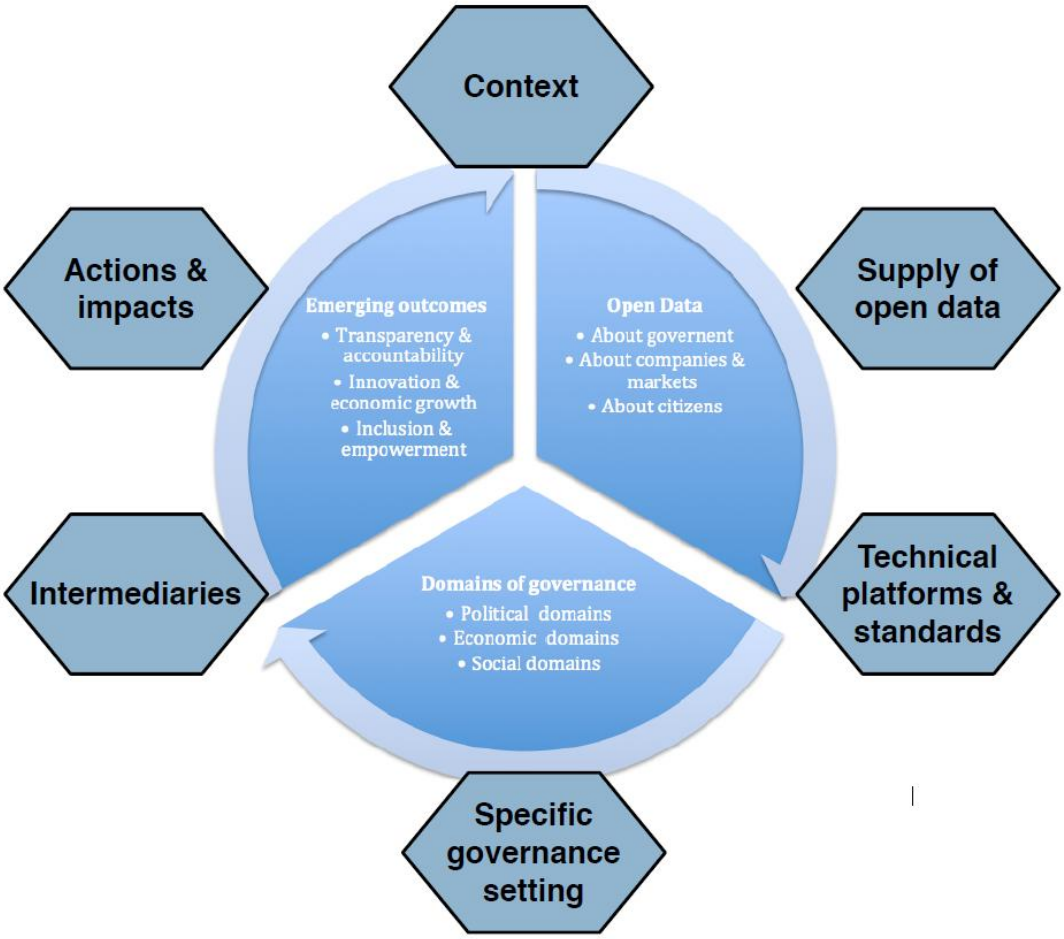

Figure 7. Open Data in Developing Countries (ODDC) Conceptual Framework

The DPER/Insight Evaluation Framework also refers to a number of economic impact studies, but these were not examined in any detail nor were recommendations provided as to how to go about doing that form of assessment. It is true that there are no standard approaches in place, however, in terms of measuring economic outcomes, and related program performance

evaluation methods and indicators, much work has been done in the open data space, and an examination of some of them are listed here along with those listed in the Evaluation Framework section 3.1.2. An analysis of what exists, what is common, and what is possible for Ireland would have been useful. As would have a comparison with what has been done in other jurisdictions.

In terms of evaluating governance, the DPER/Insight Evaluation Framework refers to the following methodologies, which could potentially be useful, but it is unknown if these were used in other jurisdictions with Open Data programs as specified in the RfT:

- Control Objective for Information and Related Technology (COBIT)
- Logical Framework Approach (LFA)
- Most Significant Change Technique (MSC)

The proposed DPER/Insight Evaluation Framework is a combination of approaches organized along the lines of the Barometer, as follows:

- ***Readiness:*** is a baseline assessment of what is in place. The report suggests that the *State of Irish Practice* sections of Chapters 5-16 of the Best Practice Handbook are the readiness assessment.
- ***Implementation:*** is to be measured against the objectives listed in the Roadmap and the actions listed in the Best Practice Handbook combined with the adoption of either the COBIT or the LAF frameworks. It was suggested that the proposed Steering and Implementation Group (SIG) be responsible for its management. Also, the OGP Independent Review Mechanism, the Open Data Barometer and the Open Knowledge Foundation Data Index can be external assessment measures.
- ***Impact:*** That a macro and micro evaluation be carried out by using the MSC technique combined with case studies to be conducted biannually and individuals from academia, civil society and research institutions can assist with that process.

The proposed approach is very pragmatic and based on the assumption that the Roadmap is the right one, that the Best Practice Handbook readiness work is complete, and that the components of the 'ecosystem' articulated in the Handbook are comprehensive, and that the recommended Open Data Board (ODB) and Steering Implementation Group (SIG) is a suitable governance strategy in the absence of a dedicated budget, business plan, human resources and an actual unit somewhere within the public sector, with administrative powers and a mandate to implement the work as listed in the Roadmap.

The ideas in the Evaluation Framework are interesting, but are not backed up with any comparative assessment of other frameworks in other jurisdictions as stipulated in the RfT, nor is justification or a critical assessment provided for this approach versus any other. There is also no explanation as to why a research framework created for the purpose of evaluating overseas development open data projects is a suitable one for Ireland. There was also no discussion about integrating this framework within current performance and evaluation frameworks within the Irish public sector, or auditing frameworks, or those commonly adopted and reported on in

other countries that have well established open data programs such as Canada, the US and the UK.

The Evaluation Framework was a review of what is most commonly known and popular in the open data field. This makes sense given the amount to time to produce this document, the fact that DPER did not specify in the RfT that the successful bidder should have an evaluation specialist in its team, nor that there be an examination of what is currently in practice within the Irish public sector. There would also need to be a way to engage the public, civil society and the private sector in this process, both in the creation of the evaluation framework and in assessing the program. Finally, there needs to be some sort of economic, social and environmental impact assessment process developed. Open data/open government are not just about innovation and commercialization as discussed earlier, and it is unfortunate that an entire section of the Roadmap is dedicated to that activity, Section 4.14 *Encouraging Commercial Use of Data* and Chapter 15 *Best Practice Standards for Encouraging Economic Reuse* in the Handbook.

The recommendations and actions are fine, however, the procurement of applications and services should be considered in a broader information strategy, while the release and use of data should not simply be for commercialization, in fact, an overemphasis on those areas may lead to false expectations. It is understood that business cases have to be made to support new strategies; however, if the view is re-oriented toward information management and good record keeping, with open data as one aspect of it, there will be internal efficiencies and growth in use. The focus should be on good administration and good governance in the public interest, and open data as an outcome of that as part of a bigger administrative program.

The pragmatic approach and some of the ideas in the Evaluation Framework report are useful, but substantively much is missing, the recommendations are based on many assumptions, and much more technical work is required, and that work should be done by a public sector evaluation and performance specialist or IT program auditors within or external to the public sector. Until that time, the proposed external evaluation methods such as Barometer indicators, the OKF Index and the OGP IRM, and others listed Chapter 16 of the Handbook combined with other indicators of openness will have to be relied upon. The Roadmap would also have to be assessed against a mission, vision and mandate before it could be accepted as valid and before being used as something against which the plan can be evaluated.

Recommendations:

- Assess current performance and evaluation frameworks within the Irish public sector, including auditing frameworks, or those commonly adopted and reported on in other countries that have well established Open Data programs such as Canada, the US and the UK and as per the RfT.
- Reassess the Open Data Barometer evaluation recommendation in the DPER/Insight report in light of its objectives and its target use and determine if it is a suitable model for a western developed national Open Data program.
- Consider high impact datasets, those of public, social and environmental significance along with those considered to be of high value

Final Remarks

Insight did much work in a short amount of time with a very small team. The time allotted to answer this CfT was too short for the number of deliverables and the requirements were very narrow. There were many requirements each calling for wide range of expertise, not all of which were present in this team (e.g., evaluation). The Roadmap and the Best Practice Handbook are good resources for the DPER to get started. The Evaluation Framework and the Publication Handbook, however, need to be strengthened. The portal was not discussed in detail in this document, but it too is a good beta or pilot version that includes many useful datasets and the platform is robust and flexible enough that metadata can be improved and discovery tools such as thesauri can be added.

The next step is for the DPER to follow the Roadmap and to get the ODB and the SIG formed, appoint an open data officer and create an open data institutional entity that is well resourced in order to implement this work. There are also many excellent human resources and well developed open data initiatives in Ireland that the government can rely upon and extend. It is hoped that the observations and recommendations provided here can be taken into consideration as the open data strategy as it moves forward.

A number of priorities were discussed here, such as good governance, records management, legal frameworks, infrastructure, organizational structures, evaluation, and public engagement to name some. The discussion in this working paper combined with those found in the DPER/Insight documents, and produced by those who responded to the call for input, should provide much food for thought. These qualify the Roadmap and should not be used as excuses to slow down the work, but should instead be used to inform it moving forward.

It is firmly believed that there is much public, civil society and private sector support and that people are willing and able to work with the DPER, the new open data office and the government entity that will be created to spearhead it. The Programmable City, AIRO, the National Centre for Geocomputation and the National Institute for Spatial and Regional Analysis, have much collective experience and expertise in data analytics, mapping, smart cities, IoT, open data, big data and critical data studies, that could be mobilized to help advance these open data plans. We look forward to working with you and hope that this Working Paper is of use.

Acknowledgements

This working paper was produced as part of the Programmable City project which is funded by the European Research Council and Science Foundation of Ireland.

Key words

open data, DPER, open government, Republic of Ireland, information management, life-cycle management, governance, law, policy, regulation, portals, recommendations, data dissemination, data discovery, evaluation, public engagement, spatial data infrastructure