

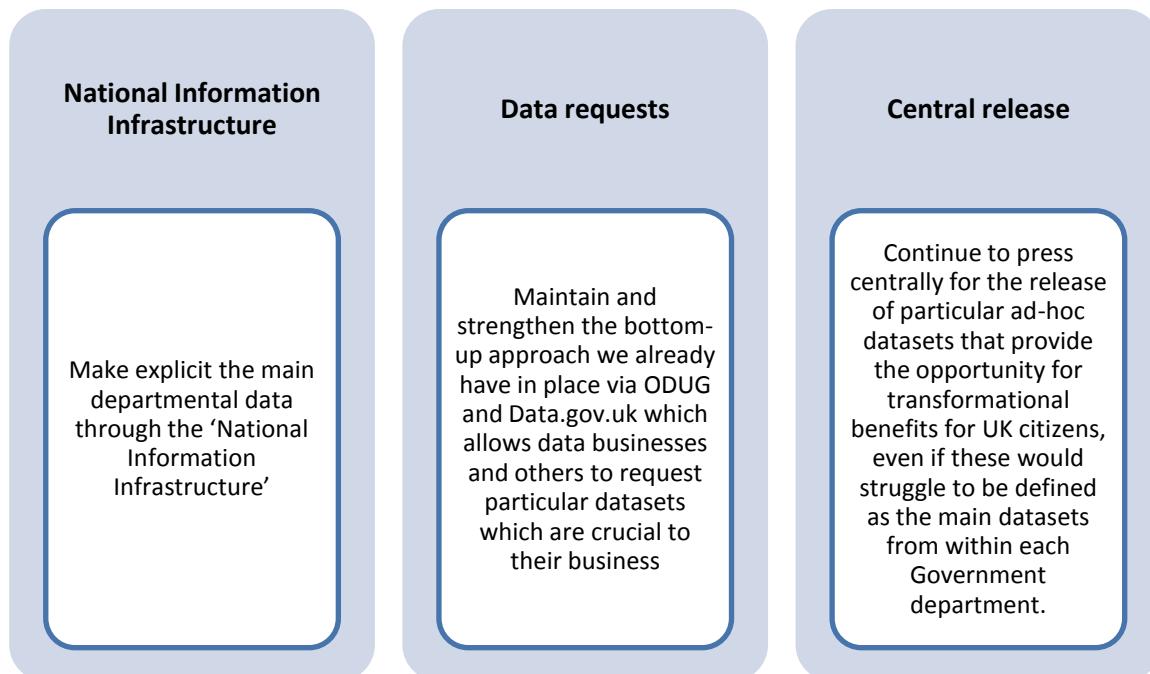
Guidance to Departments on Developing the UK National Information Infrastructure

Cabinet Office, July 2013

Version 4.0

Background

In response to recommendations in the Shakespeare Review of Public Sector Information, the government set out a three pronged approach to releasing further open data:



To support this work it committed to “set out the data we already have by department and invite comment from business users about what data they would like released”. This commitment to developing an inventory of the data held by Government reflects existing policy and builds on good practice in fulfilling the regulations on Public Sector Information. For instance in terms of the Transparency agenda:

- a partial inventory of the data that central government departments, as well as many Arms Length Bodies and Local Authorities, hold is currently already available on Data.gov.uk
- the Open Data White Paper¹ committed government (public bodies) to identifying and publishing priority datasets and over time removing barriers to the publication of more datasets.
- through Open Data Strategies, Departments have set out a clear timeline for publication for a subset of the data they hold and a route for users to request further data to be made open².

In terms of the broader landscape of PSI, Departments will likely still have information they are using following the HMT led Data Review from 2011 and guidance from OPSI on information asset registers. Departments are already likely to have an inventory if they are following the CESG Information Asset Maturity Model³.

¹ Cabinet Office (28 June 2012) [Open Data White Paper: Unleashing the Potential](#) A document that seeks to lay out what citizens, businesses and the public sector can expect from government to help unlock the benefits of Open Data.

² [Departmental Open Data Strategies](#)

³ IA Maturity Model <http://www.cesg.gov.uk/policyguidance/IAMM/Pages/index.aspx>

Given the breadth of stakeholders involved the Cabinet Office Transparency Team is setting out a collaborative process for identifying those datasets which should be part of the 'National Information Infrastructure'. This document sets out this guidance as well as the steps towards developing a first draft. This document will be circulated in draft to departments in the first instance and then subsequently published publicly.

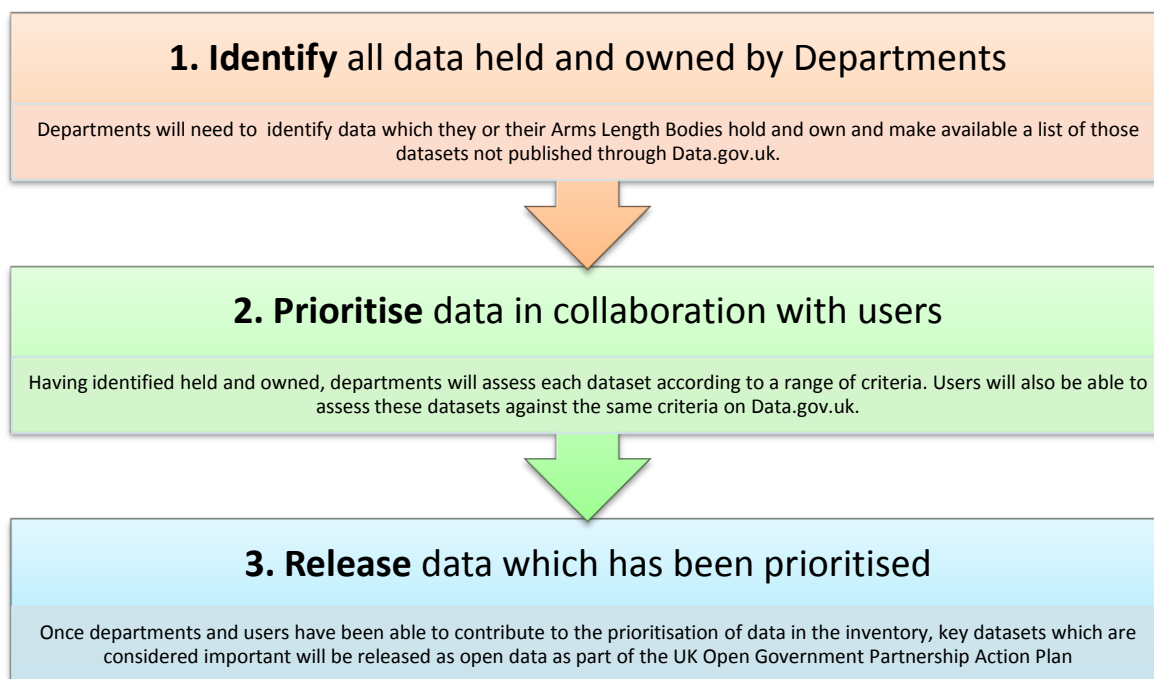
Departments will also revisit their Open Data Strategies as part of this process. Given the development of a comprehensive inventory of the data held by departments and an associated timetable for release of this information, departments may decide to refocus their Open Data Strategy so that it highlights the activity which they are undertaking to support their data release work.

Rationale for Exercise

The UK Government has made the commitment to move towards a "presumption to publish" public sector information. There have been positive movements towards this by a number of departments. The recent Shakespeare Review of Public Sector Information highlighted a number of areas where the transparency and open data agenda could be strengthened and this included a recommendation to identify "National Core Reference Data". The Government Response to the review responded to this by setting out plans to develop a National Information Infrastructure, which would encompass both data which could be described as "core reference data" but also other important departmental data. In order to identify this National Information Infrastructure, departments will need in the first instance to publish a list of all the data they hold as an inventory.

Overview of the process

We are asking Departments to undertake three phases of activity in the first instance which will end up in identifying the National Information Infrastructure and make more data available as open data. Broadly these stages can be characterised as **'Identify'**, **'Prioritise'** and **'Release'**.



Outputs of exercise

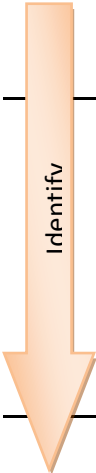
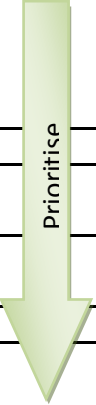
There will be a number of outputs from this exercise. By October 2013, this process will have created:

- An inventory of government data on Data.gov.uk
- Updated Departmental Open Data Strategies
- Prioritised datasets for release by departments

On an ongoing basis, Data.gov.uk will be maintained by departments as an inventory of government data.

Timetable and phasing

The initial work to create a first cut of an inventory on Data.gov.uk needs to be completed in time for the Open Government Partnership conference at the end of October 2013. From then on departments and their ALBs should continue to maintain the records which they have made available on Data.gov.uk so that it can become a truly comprehensive inventory of government owned and held data, as well as a catalogue of the data published by Government.

	From now	Depts and ALBs should update Data.gov.uk to ensure that all data they publish is recorded on Data.gov.uk. Guidance on how to do this is available here .
	w/c 15th July	Transparency Team will make available to departmental contacts: <ul style="list-style-type: none"> • data dumps of records currently available through Data.gov.uk • a template for the spreadsheet to upload unpublished data • where relevant a copy of the HMT Data Review spreadsheets • outstanding data requests
	From 5th August	Start to upload unpublished data with new tranches added every Friday until 30 th August Launch of functionality on Data.gov.uk which will allow assessment of all datasets against NII criteria.
	By 30th August	Complete uploading of unpublished data to Data.gov.uk
	By 2nd September	Departments to complete assessment of datasets against criteria on Data.gov.uk
	By 16th September	Complete analysis of datasets within National Information Infrastructure
	23rd September	Report to Public Sector Transparency Board
	7th October	Start Cabinet Committee Clearance for proposed National Information Infrastructure
	31st October	Publication of the National Information Infrastructure in the UK OGP Action Plan

Stage One - Identify

Departments and their Arms Length Bodies will already keep details of the data they hold, including datasets which are made available commercially at present and maintain⁴ and much of this will already be published on Data.gov.uk. The Cabinet Office Transparency Team will support Departments to identify unpublished dataset by providing them with the following:

- a list of the datasets which are already published on Data.gov.uk;
- (where relevant) a copy of the Department's return for the HMT led Data Review from 2011;
- outstanding data requests which have been submitted via the data request form on Data.gov.uk and remain unfulfilled.

The Cabinet Office will supply a template spreadsheet to departments in order to upload unpublished datasets to Data.gov.uk starting **on 5th August**. Once initial data has been started to be uploaded by departments Cabinet Office Transparency Team will arrange a meeting to discuss further work needed. Departments will need consider how they would like to ensure their Ministers are involved in the process as well as how they would like to involve their Transparency Sector Board if they have one. Departments will need to have completed uploading of their unpublished data **by 30th August**.

If departments or ALBs find that data is published but not yet available on Data.gov.uk, they should make arrangements for it to be part of the catalogue on Data.gov.uk. Guidance on how to publish on Data.gov.uk is available [here](#). The "live" inventory on Data.gov.uk will be able to be updated and amended by departments with additional datasets to as they are collected.

Scope of Data

The definition of a dataset in the inventory will follow that used in the Datasets Code of Practice⁵ i.e.: information comprising a collection of information held in electronic form where all or most of the information in the collection

- a. has been obtained or recorded for the purpose of providing a public authority with information in connection with the provision of a service by the authority or the carrying out of any other function of the authority,
- b. is factual information which
 - i. is not the product of analysis or interpretation other than calculation, and
 - ii. is not an official statistic (within the meaning given by section 6(1) of the Statistics and Registration Service Act 2007), and
 - iii. remains presented in a way that (except for the purpose of forming part of the collection) has not been organised, adapted or otherwise materially altered since it was obtained or recorded

Scope of Organisations

Central government departments and their Arms Length Bodies, including Trading Funds, will be included in this exercise. All data which these organisations hold

Stage two - Prioritise

⁴ Best practice in [The re-use of Public Sector Information: A guide to the regulations and best practice](#) produced by The National Archives sets out that departments should keep Information Asset Lists of information that is made available by Departments.

⁵ <http://www.data.gov.uk/consultation/code-of-practice>

Having begun to publish their data inventory through Data.gov.uk, departments should then start to assess the importance of the data they hold so that a decision can be made whether it would be included in the National Information Infrastructure or not.

The Cabinet Office Transparency Team will develop functionality on Data.gov.uk so that departments can assess data at the same time as users, in an open manner. We are currently developing functionality on Data.gov.uk so that departments will be able to make these assessments directly on the website and once this has been developed we will circulate update guidance setting out the practicalities of this. Departments will need to complete assessments of the importance of the data in their inventory **by 16th September**.

Criteria for Assessment

In the [Government Response to the Shakespeare Review](#) we set out criteria by which departments and users should evaluate data in the inventory. These criteria are:

- Economic Growth
- Social Growth
- Effective public services
- Connective reference data
- Other Key data

A. Economic Growth – input to be lead by users

Data will be assessed against whether it would support economic growth if it was made available openly. This assessment will be led by contributions by potential users of the data giving details of how the data could support their commercial activity. This is similar to the way in which users at the moment set out the likely business benefit of data through the Data.gov.uk data request process. Input for this criterion will be collected from users qualitatively (i.e. they will be asked to elaborate on the potential economic impact). Users will be asked to consider the following questions in their input on this criterion:

- If open, could it stimulate growth in the UK economy?
- Is it being requested by business?
- Would it enable more efficient functioning of markets and reduce the cost of living for citizens?

B. Social Growth – input to be lead by users

Data will be assessed against whether it would support social growth if it was made available openly. This assessment will be lead by contributions by potential users of the data giving details of how the data could support activity such as social action. Input for this criterion will be collected from users qualitatively (i.e. they will be asked to elaborate on the potential social growth impact). Users will be asked to consider the following questions in their input on this criterion:

- Is it requested by campaigning groups?
- If open, would it help stimulate volunteering and self-help?
- Could it aid in promoting social development and change?

C. Effective public services – input to be jointly lead by departments and users

Data will be assessed against whether, if it was made available openly, it would support effective public services. This assessment will be jointly lead by contributions from departments and potential users. Input for this criterion will be collected from users qualitatively (i.e. they will be asked to elaborate on the potential impact on public services). Departments and users will be asked to consider the following questions in their input on this criterion:

- Is the data fundamental to the operation of each department?
- Could it be used to hold government to account if released openly?
- If open, could it aid the efficiency of public services and the running of government?
- Could it aid the public in making choices about which public services to use if opened up?

D. Connective reference data – input to be lead by UK Linked Data Working Group

Data will be assessed by the UK Linked Data Working Group (UKGovLD) Delivery Team to ascertain whether it represents connective reference data. In making this assessment they will consider:

- Whether data is commonly referenced by other datasets or applications
- Data that is most valuable for interconnecting different datasets and making different datasets comparable
- Data that is durable and persistent (needs to be relied upon to still be there over time)
- Whether the government is the sole owner of this information, or is uniquely well placed to provide the data.

The UKGovLD will then consult with the wider community on prioritisations against their assessments.

E. Other key data

Departments and users will also be able to highlight other potential beneficial outcomes of releasing the data openly.

Overall Assessments

Once initial assessments have been undertaken on data in each inventory by departments and users, departments along with the Cabinet Office will review these and prioritise the datasets on the basis of responses to each of the individual criteria. In relation to criteria A to C the prioritisation of datasets will be based on:

- the strength of the argument put against each criterion;
- the number of users who highlight an impact against each criterion;
- the number of different impacts highlighted under each criterion; and
- the potential number of users who could exploit the data if released openly.

This will be converted into an overall assessment for each dataset. In terms of the release phase of the process this guidance will be updated in August setting out how Departments should go about identifying those datasets within the National Information Infrastructure to prioritise for early release.

FAQs

1. What types of datasets will be included in the inventory?

The inventory will include datasets which describe information across a range of subject domains including:

- Corporate
- Delivery
- Estates
- Finance
- HR
- Policy
- Research

2. Should we include aggregate, derived or combined information which we don't yet publish?

At this point we are not interested in including in the inventory aggregate, combined or derived data which departments don't at this point in time publish. Departments will already have a programme for publishing aggregate information in the form of statistics and many of these should be available through Data.gov.uk.

3. What if the datasets include personal or security information?

The inventory will include all data which is held by departments and ALBs and this will include datasets which hold personal or security information. Regardless of whether it would be possible for the department to publish some or all of the data in these datasets these should still be included in the inventory.

4. How does this process build on current work?

This process also builds on best practice which has been developed with the Open Data User Group and others. For instance it will incorporate:

- Departments being clear to users what data they hold. Previously users have had to make requests for data which may or may not be held by Government and may not be held by the department they think it is.
- Clear transparency about process. All inventories will be published on Data.gov.uk and their assessment will be carried out in the open, by both departments and users at the same time.
- Builds on the Data.gov.uk platform. Functionality to support this process of making departmental inventories public will be incorporated into the record structure of Data.gov.uk so that it becomes a catalogue of **all data held** by departments not just that which is currently published.

5. How does this process work with the current 'Data Request' process on Data.gov.uk?

There is currently functionality on Data.gov.uk where users can complete a form making a "data request" for the release of a dataset. This process was developed in conjunction with the Open Data User Group (ODUG) and helped support their building of cases to put the Data Strategy Board (DSB). Following the Government Response to the Shakespeare Review, changes to the DSB and the commitment to provide a National Information Infrastructure this process is being extended. This will further develop the functionality of Data.gov.uk to support engagement on the prioritisation process for the National Information Infrastructure and the release of datasets as open data.

Departments should carry on pursuing these requests in parallel with identifying data for the inventory and assessing them against NII criteria.

6. Will departments still need to complete the Quarterly Written Ministerial Statement for Transparency?

Departments will continue to complete the Quarterly Written Ministerial Statement in the short term. The data.gov.uk team are currently developing the necessary tools to measure and compare departments' adherence to commitments. Once these are complete, data.gov.uk will automatically produce reports and these will be made public. When this functionality has been bedded in, the WMS returns will be stopped. The aim is that the last return will be for Q3 2013/14, published in March 2014. Before then, departments will still be expected to submit returns on a quarterly basis.

Period	Return due	QWMS to be published
Q1 2013 / 14	06-Sep-2013	Before end Sept 13
Q2 2013 / 14	06-Dec-2013	Before end Dec 13
Q3 2013 / 14	07-Mar-2013	Before end Mar 13

7. Will departments have to do anything on the UK Open Government Partnership UK National Action Plan during this period as well?

The UK is the lead co-chair for the Open Government Partnership until November 2013. There is a global movement to increase transparency and it is growing as more countries join the Open Government Partnership.

The final Open Government Partnership 2013 UK National Action Plan, to be published at the OGP annual summit at the end of October, will set the direction for the UK Government's plans on open policymaking, transparency and releasing open data. The draft National Action Plan is now out for consultation (closes 19 September). Views on the commitments contained within the plan are currently being sought from civil society organisations and these will inform the development of the final plan. A team in the Cabinet Office Transparency Team will be in contact with Departments to discuss the ideas currently highlighted in the draft plan as well as any potential commitments which could be included by departments in the final draft in October.

8. How does this work take into account the recent commitment by the UK to the G8 Open Data Charter?

G8 leaders signed the Open Data Charter on 18 June 2013 (see Annex B). The Open Data Charter sets out 5 strategic principles that all G8 members will act on. These include an expectation that all government data will be published openly by default, alongside principles to increase the quality, quantity and re-use of the data that is released. G8 members have also identified 14 high-value areas – from education to transport, and from health to crime and justice – from which they will release data. These will help unlock the economic potential of open data, support innovation and provide greater accountability.

As part of the Charter there is a list of dataset domains which countries have committed to ensuring data is made available as open data. Departments should consider whether they hold data within these domains and reflect this in their assessments to identify the National Information Infrastructure.

9. How does this process relate to the Right to Data?

Before Parliament's Summer Recess, which starts on 18 July, we will commence amendments to the Freedom of Information Act which will give a 'right to data' wherever release of a data set is required by that Act. Where reasonably practicable, any public authority will be required to release such a data set in a machine readable format. Commencement will be accompanied by a data sets code of practice issued under the FOI Act that will set out best practice for public authorities when complying with the new duty of FOI.

The inventory which will be built through this process will use the same definition of datasets as set out in the changes to the FOI act.

10. How does this process relate to the changes to the EU PSI Directive?

The EU Directive on the re-use of PSI will be adopted in Summer 2013. The Government aims to transpose the terms of the revised Directive into UK legislation during 2014-15.⁶ The Directive and corresponding UK legislation will give an absolute right to most public sector data being made available for re-use, except in limited circumstances where privacy, legal or charging exceptions apply. It will establish marginal cost pricing as the default position, which is in relation to the provision of data in a machine readable format at mostly zero cost. It reinforces oversight and a mandatory framework for PSI re-use. Combined with the main data sets provision set out above, this will give a strong legal underpinning to the availability of this data for re-use. We aim for all the main data to be freed no later than 2015.

11. How does this process relate to the EU INSPIRE Directive and the UK INSPIRE Regulations?

The EU INSPIRE Directive was incorporated into UK law in 2009 and sets a timetable for public authorities to publish spatial data held as part of their public task. INSPIRE also requires Network (web data) Services to be in place to enable and support this data publishing. The next INSPIRE milestone is 3 December 2013 by when data in scope of INSPIRE must have metadata registered at data.gov.uk and organisations with INSPIRE datasets must have INSPIRE View Services and Download Services to support the use of the data.

Under INSPIRE conditions can be placed on the access or use of data sets and Network Services. These will usually be by a licence but can include charges for access or use of data set or data service. The Government Licensing Framework was developed to be used for INSPIRE datasets. The guidance has always been to use OGL in preference to any other licence. INSPIRE does not allow for charges to be made to view metadata. There are limited circumstances in which metadata may be restricted where defence, security, public safety or copyright and confidentiality issues apply. Similarly, limits may be placed on access to and use of data and services.

While INSPIRE Annex III metadata is not due to be available until 3 December 2013 it does not affect the requirement to list data in the inventory of the UK National Information Infrastructure. A more detailed metadata record must still be produced for the December deadline and the UKNII inventory process will help any organisations which have not yet completed a data audit or catalogued the data they hold prepare for the INSPIRE deadline.

12. How does this process relate to Publication Schemes?

⁶ The UK implemented a European Directive on the re-use of Public Sector Information in 2005. Since then the UK has made considerable progress in promoting and encouraging re-use. Indeed, in recent European Commission funded reviews the UK is shown as being the leading member state on re-use.

The Freedom of Information Act requires every public authority to have a publication scheme⁷, approved by the Information Commissioner's Office (ICO), and to publish information covered by the scheme. The scheme must set out your commitment to make certain classes of information routinely available, such as policies and procedures, minutes of meetings, annual reports and financial information. To help the ICO has developed a model⁸ publication scheme. The main class of interest to the NII activity are:

- **'Lists and Registers':** "Information held in registers required by law and other lists and registers relating to the functions of the authority."
- **'Who we are and what we do.':** "Organisational information, locations and contacts, constitutional and legal governance."
- **'What we spend and how we spend it.':** "Financial information relating to projected and actual income and expenditure, tendering, procurement and contracts."
- **'What our priorities are and how we are doing.':** "Strategy and performance information, plans, assessments, inspections and reviews."

The scope of the NII is broader than this, but we would expect datasets referenced within the Publication Scheme to be listed in the inventory.

⁷http://www.ico.org.uk/for_organisations/freedom_of_information/guide/publication_scheme?hidecookiesbanner=true#what-information-do-we-need-to-publish-6

⁸http://www.ico.org.uk/for_organisations/freedom_of_information/guide/~media/documents/library/Freedom_of_Information/Detailed_specialist_guides/generic_scheme_v1.ashx

Annex A – Data fields to be completed for datasets not yet published

Field Name	Field Information	Future use of this information	Input format expected
Title	The title of the item.	So users can easily reference and understand the item.	Free text in Plain English.
Description	The main description of the item.	So users can see details of the item and judge its worth to them.	Free text in Plain English. The first line should be a brief description, followed by more if needed.
Owner	The organisation directly credited with or associated with the item.	So users can identify the item's owner.	As stated on Data.gov.uk publication hierarchy
Date to be published	If there is already a public commitment to release this, when the item will be released,	So users can know when to expect the data to be released.	Planned Date of release: in the YYYYMMDD format
Release notes	Extra information on the release of the data. This might include reasons why the data will not be released.	So users are aware of the reasoning for non-release of data.	Free text in plain English

Annex B - G8 Open Data Charter and Technical Annex

Published 18 June 2013

Contents

1. [Principle 1: Open Data by Default](#)
2. [Principle 2: Quality and Quantity](#)
3. [Principle 3: Usable by All](#)
4. [Principle 4: Releasing Data for Improved Governance](#)
5. [Principle 5: Releasing Data for Innovation](#)
6. [Technical annex](#)

Preamble

1) The world is witnessing the growth of a global movement facilitated by technology and social media and fuelled by information – one that contains enormous potential to create more accountable, efficient, responsive, and effective governments and businesses, and to spur economic growth.

Open data sit at the heart of this global movement.

2) Access to data allows individuals and organisations to develop new insights and innovations that can improve the lives of others and help to improve the flow of information within and between countries. While governments and businesses collect a wide range of data, they do not always share these data in ways that are easily discoverable, useable, or understandable by the public.

This is a missed opportunity.

3) Today, people expect to be able to access information and services electronically when and how they want. Increasingly, this is true of government data as well. We have arrived at a tipping point, heralding a new era in which people can use open data to generate insights, ideas, and services to create a better world for all.

4) Open data can increase transparency about what government and business are doing. Open data also increase awareness about how countries' natural resources are used, how extractives revenues are spent, and how land is transacted and managed. All of which promotes accountability and good governance, enhances public debate, and helps to combat corruption. Transparent data on G8 development assistance are also essential for accountability.

5). Providing access to government data can empower individuals, the media, civil society, and business to fuel better outcomes in public services such as health, education, public safety, environmental protection, and governance. Open data can do this by:

- showing how and where public money is spent, providing strong incentives for that money to be used most effectively;
- enabling people to make better informed choices about the services they receive and the standards they should expect.

6) Freely-available government data can be used in innovative ways to create useful tools and products that help people navigate modern life more easily. Used in this way, open data are a catalyst for innovation in the private sector, supporting the creation of new markets, businesses, and jobs. Beyond government, these benefits can multiply as more businesses adopt open data practices modelled by government and share their own data with the public.

7) We, the G8, agree that open data are an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of our citizens and allow innovation and prosperity to flourish.

8) We therefore agree to follow a set of principles that will be the foundation for access to, and the release and re-use of, data made available by G8 governments. They are:

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

9) While working within our national political and legal frameworks, we will implement these principles in accordance with the technical best practises and timeframes set out in our national action plans. G8 members will, by the end of this year, develop action plans, with a view to implementation of the Charter and technical annex by the end of 2015 at the latest. We will review progress at our next meeting in 2014.

10) We also recognise the benefits of open data can and should be enjoyed by citizens of all nations. In the spirit of openness we offer this Open Data Charter for consideration by other countries, multinational organisations and initiatives.

1. Principle 1: Open Data by Default

11) We recognise that free access to, and subsequent re-use of, open data are of significant value to society and the economy.

12) We agree to orient our governments towards open data by default.

13) We recognise that the term government data is meant in the widest sense possible. This could apply to data owned by national, federal, local, or international government bodies, or by the wider public sector.

14) We recognise that there is national and international legislation, in particular pertaining to intellectual property, personally-identifiable and sensitive information, which must be observed.

15) We will:

- establish an expectation that all government data be published openly by default, as outlined in this Charter, while recognising that there are legitimate reasons why some data cannot be released.

2. Principle 2: Quality and Quantity

16) We recognise that governments and the public sector hold vast amounts of information that may be of interest to citizens.

17) We also recognise that it may take time to prepare high-quality data, and the importance of consulting with each other and with national, and wider, open data users to identify which data to prioritise for release or improvement.

18) We will:

- release high-quality open data that are timely, comprehensive, and accurate. To the extent possible, data will be in their original, unmodified form and at the finest level of granularity available;
- ensure that information in the data is written in plain, clear language, so that it can be understood by all, though this Charter does not require translation into other languages;
- make sure that data are fully described, so that consumers have sufficient information to understand their strengths, weaknesses, analytical limitations, and security requirements, as well as how to process the data; and
- release data as early as possible, allow users to provide feedback, and then continue to make revisions to ensure the highest standards of open data quality are met.

3. Principle 3: Usable by All

19) We agree to release data in a way that helps all people to obtain and re-use it.

20) We recognise that open data should be available free of charge in order to encourage their most widespread use.

21) We agree that when open data are released, it should be done without bureaucratic or administrative barriers, such as registration requirements, which can deter people from accessing the data.

22) We will:

- release data in open formats wherever possible, ensuring that the data are available to the widest range of users for the widest range of purposes; and
- release as much data as possible, and where it is not possible to offer free access at present, promote the benefits and encourage the allowance of free access to data. In many cases this will include providing data in multiple formats, so that they can be processed by computers and understood by people.

4. Principle 4: Releasing Data for Improved Governance

23) We recognise that the release of open data strengthens our democratic institutions and encourages better policy-making to meet the needs of our citizens. This is true not only in our own countries but across the world.

24) We also recognise that interest in open data is growing in other multilateral organisations and initiatives.

25) We will:

- share technical expertise and experience with each other and with other countries across the world so that everyone can reap the benefits of open data; and
- be transparent about our own data collection, standards, and publishing processes, by documenting all of these related processes online.

5. Principle 5: Releasing Data for Innovation

26) Recognising the importance of diversity in stimulating creativity and innovation, we agree that the more people and organisations that use our data, the greater the social and economic benefits that will be generated. This is true for both commercial and non-commercial uses.

27) We will:

- work to increase open data literacy and encourage people, such as developers of applications and civil society organisations that work in the field of open data promotion, to unlock the value of open data;
- empower a future generation of data innovators by providing data in machine-readable formats.

6. Technical annex

1) We, the G8, have consulted with technical experts to identify some best practices (part one) and collective actions (part two) that we will use to meet the principles set out in the G8 Open Data Charter.

2) While working within our national political and legal frameworks, we agree to implement these practices as quickly as possible and aim to complete our activities by 2015 at the latest. This will be done in accordance with the timeframes in our national action plans.

3) The Annex constitutes a 'living' set of guidelines that may be subject to amendments after consideration of emerging technology solutions or practical experience gained during the course of implementation of the G8 Open Data Charter.

6.1 Part One - Best Practices

Principle 1: Open Data by Default

4) We recognise the importance of open data and we will establish an expectation that all government data be published openly by default.

5) We will:

- define our open data position in a public statement of intent, such as an announcement, strategy or policy, so that our plans for progressing the open data agenda in our jurisdictions are clear;
- publish a national action plan to provide more specific details on our plans to release data according to the principles in the G8 Open Data Charter; and
- publish data on a national portal so that all government data that has been released can be found easily in one place. A portal may be a central website from which data can be downloaded, or a website which lists all open government data stored at a different location. Each portal will include a registry file that lists all the data and metadata used on the portal, as well as providing APIs for developers. Where it is yet not possible to publish all data on a portal, the location of data will be communicated clearly and not moved without notice.

Principle 2: Quality and Quantity

6) We commit to releasing data that are both high in quality as well as high in quantity. When releasing data, we aim to do so in a way that helps people to use and understand them. This will help to increase the interoperability of data from different policy areas, businesses or countries.

7) We will:

- use robust and consistent metadata (i.e. the fields or elements that describe the actual data);
- publish and maintain an up-to-date mapping of the core descriptive metadata fields across G8 members to enable easier use and comprehension by people from around the world. This will allow countries, in the G8 and beyond, who do not currently have a data portal to consider adopting the metadata fields included in this mapping;
- ensure data are fully described, as appropriate, to help users to fully understand the data. This may include:
 - Documentation that provides explanations about the data fields used;
 - Data dictionaries to link different data; and
 - A user's guide that describes the purpose of the collection, the target audience, the characteristics of the sample, and the method of data collection.
- listen to feedback from data users to improve the breadth, quality and accessibility of data we offer. This could be in the form of a public consultation on the national data strategy or policy, discussions with civil society, creation of a feedback mechanism on the data portal, or through other appropriate mechanisms.

Principle 3: Usable by All

8) We agree to release data in a way that helps all people find and re-use them.

9) We will:

- make data available in convenient open formats to ensure files can be easily retrieved, downloaded, indexed, and searched by all commonly used Web search applications. Open formats, for example non-proprietary CSV files, are ones where the specification for the format is available to anyone for free, thereby allowing the data contained in a file to be opened by different software programmes.

Principle 4: Releasing Data for Improved Governance

10) We recognise that data are a powerful tool to help drive government effectiveness, efficiency and responsiveness to citizen needs while fuelling further demand for open data.

11) We will:

- develop links with civil society organisations and individuals to allow the public to provide feedback on the most important data they would like released;
- be open about our own data standards, so that we take into account:
 - Data that are released by other national and international organisations
 - The standards emerging from other international transparency initiatives; and
- document our own experiences of working with open data by, for example, publishing technical information about our open data policies, practices, and portals so that the benefits of open data can be enjoyed in other countries.

Principle 5: Releasing Data for Innovation

12) We agree that our citizens can use our data to fuel innovation in our own countries and around the world. We recognise that free access to, and reuse of, open government data are an essential part of this.

13) We will:

- support the release of data using open licences or other relevant instruments - while respecting intellectual property rights - so that no restrictions or charges are placed on the re-use of the information for non-commercial or commercial purposes, save for exceptional circumstances;
- ensure data are machine readable in bulk by providing data that are well structured to allow automated processing and access with the minimum number of file downloads;
- release data using application programming interfaces (APIs), where appropriate, to ensure easy access to the most regularly updated and accessed data; and
- encourage innovative uses of our data through the organisation of challenges, prizes or mentoring for data users in our individual jurisdictions.

6.2 Part Two - Collective Actions

Action 1: G8 National Action Plans

- We will publish individual action plans detailing how we will implement the Open Data Charter according to our national frameworks (October 2013)
- We will report progress on an annual basis (via the G8 Accountability Working Group) (2014 and 2015)

Action 2: Release of high value data

- We recognise the following as areas of high value, both for improving our democracies and encouraging innovative re-use of data.

Data Category* (alphabetical order)

Example datasets

Companies	Company/business register
Crime and Justice	Crime statistics, safety
Earth observation	Meteorological/weather, agriculture, forestry, fishing, and hunting

Data Category* (alphabetical order)	Example datasets
Education	List of schools; performance of schools, digital skills
Energy and Environment	Pollution levels, energy consumption
Finance and contracts	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent)
Geospatial	Topography, postcodes, national maps, local maps
Global Development	Aid, food security, extractives, land
Government Accountability and Democracy	Government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/gifts
Health	Prescription data, performance data
Science and Research	Genome data, research and educational activity, experiment results
Statistics	National Statistics, Census, infrastructure, wealth, skills
Social mobility and welfare	Housing, health insurance and unemployment benefits
Transport and Infrastructure	Public transport timetables, access points broadband penetration

- In accordance with the principles of “open by default” and “quality and quantity” we will work towards the progressive publication of these data.
- As a first step, we will collectively make key datasets on National Statistics, National Maps, National Elections and National Budgets available and discoverable (from June 2013), and we will work towards improving their granularity and accessibility (by December 2013)
- We recognise that collective action by all G8 members has the potential to unlock barriers and foster innovative solutions to some of the challenges we are facing. We therefore agree on a mutual effort to increase the supply of open government data available on key functions of our States, such as democracy and environment. We will work on identifying datasets in these areas by December 2013, with an aim to release them by December 2014.
- We will set out in our national action plans how and when we will release data under the remaining categories according to our national frameworks (October 2013).

Action 3: Metadata mapping

- We have contributed to and commit to maintaining the G8 metadata mapping exercise (June 2013)
- This mapping can be viewed on Github and comprises a [collective mapping ‘index’](#) across G8 member’s metadata, and a [detailed page](#) on each G8 member use of metadata within their national portal.

*categories and datasets to be finalised by December 2015