

Open Data Barometer: 2013 Researchers Handbook

This is an archived copy of the 2013 Open Data Barometer Researchers Handbook. You can find out more about the Open Data Barometer at www.opendatabarometer.org

The Barometer study was undertaken with support of the Open Data Institute, and the World Wide Web Foundation, and formed a pilot as part of the Common Assessment Methods components of the Exploring the Emerging Impacts of Open Data project.



WORLD WIDE WEB
FOUNDATION



This work is licensed under a [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/)

Please cite as: Web Foundation. 'Open Data Barometer: 2013 Researchers Handbook', June 2013

Contact: tim.davies@webfoundation.org

Contents:

[1.1. About this handbook](#)

[1.2. Research Methodology](#)

[1.3. Using the research platform](#)

[Researchers](#)

[Getting started](#)

[Entering data](#)

[Flagging sheets for review](#)

[Regional coordinators](#)

[Reviewers](#)

[Researchers](#)

[Reviewing and responding to comments](#)

[Final review](#)

[Support](#)

[1.4. Justifications and sources](#)

[1.5. Additional research guidance](#)

[General definitions](#)

[Context questions](#)

[C1: To what extent is there a well-resourced open government data initiative in this country?](#)

[C2: To what extent does the country have a functioning right-to-information law?](#)

[C3: To what extent is there a robust legal or regulatory framework for protection of personal data in your country?](#)

[C4: To what extent are civil society and information technology professionals engaging with the government regarding open data?](#)

[C5: To what extent are city or regional governments running their own open data initiatives?](#)

[C6: To what extent is government directly supporting a culture of innovation with open data through competitions, grants or other support?](#)

[C7: To what extent is high quality training available for individuals or businesses wishing to increase their skills or build businesses to use open data?](#)

[C8: To what extent are the following forms of training available in the country?](#)

[C9: To what extent are academic institutions in the country opening up their data?](#)

[C10: To what extent are civil society in the country opening up their own data?](#)

[C11: To what extent are businesses in the country opening up their own data?](#)

[Impact questions](#)

[I1: To what extent has open data had a noticeable impact on increasing government efficiency and effectiveness?](#)

[I2: To what extent has open data had a noticeable impact on increasing transparency and accountability in the country?](#)

[I3: To what extent has open data had a noticeable impact on environmental sustainability in the country?](#)

[I4: To what extent has open data had a noticeable impact on increasing the inclusion of](#)

[marginalised groups in policy making and accessing government services?](#)

[I5: To what extent has open data had a noticeable positive impact on the economy?](#)

[I6: To what extent are entrepreneurs successfully using open data to build new businesses in the country?](#)

[Dataset questions](#)

[Is it open? Identifying whether a dataset is open](#)

- [a\) Does the data exist?](#)
- [b\) Is it available online from government in any form?](#)
- [c\) Is the dataset provided in machine readable formats?](#)
- [d\) Is the machine readable data available in bulk?](#)
- [e\) Is the dataset available free of charge?](#)
- [f\) Is the data openly licensed?](#)
- [g\) Is the dataset up to date?](#)
- [h\) Is the publication of this dataset sustainable?](#)
- [i\) Was it easy to find information on this dataset?](#)
- [j\) Are data URIs provided for key elements in the dataset?](#)

[Dataset details](#)

- [D1: Map data \(full map coverage of the country\)](#)
- [D2: Land ownership data](#)
- [D3: Government service directory](#)
- [D4: Detailed census data](#)
- [D5: Detailed government budget](#)
- [D6: Detailed data on government spend](#)
- [D7: Company register](#)
- [D8: Legislation](#)
- [D9: Public transport timetables](#)
- [D10: International trade data](#)
- [D11: Health sector performance](#)
- [D12: Primary or secondary education performance data](#)
- [D13: Crime statistics](#)
- [D14: National environment statistics](#)
- [D15: National election results](#)

Section 1: process and guidance

1.1. About this handbook

This Handbook provides detailed guidance for researchers, regional coordinators and reviewers of the 2013 Open Data Barometer. Please read the first section in detail before starting your research or review. The detailed question guidance can be consulted as you work through the survey form.

1.2. Research Methodology

For the expert survey portion of the 2013 Open Data Barometer, a lead country researcher performs desk research and consults with key informants to score the indicators for his/her country. Those initial scores are then reviewed by the regional coordinator, who may send back a portion of those questions for improvement or correction based on the standards described in this codebook. When the regional coordinator is satisfied with the state of the indicators, an expert reviewer will be invited to review the indicators and will flag certain questions for additional research, correction or clarification. Regional coordinators will review all reviewer feedback to determine which indicators need to be returned to the lead researcher for further strengthening; this back-and-forth will lead to a final set of well-researched and well-sourced indicators.

The time period under study is the 12-month period from June 2012 to June 2013. Questions take two main forms:

- **‘To what extent’** questions, with 11 score choices from 0 to 10 with 10, 5 and 0 score choices containing detailed scoring criteria to guide the researcher in his/her selection of the most appropriate score.
- **‘Dataset questions’** broken down into 10 yes/no questions, and a number of prompts for additional evidence.

There is also one question that asks for example URLs of training providers.

For every question it is important that researchers provide details of the sources used to answer that question, and any additional information requested. This will support reviewers to check the answers given.

Most questions can be answered drawing upon online sources, published materials and desk research. Tips on sources and searching for information are provided alongside each question. In a number of cases you may wish to interview or consult with open data experts, NGOs, journalists or government officials to identify the appropriate answer to a question. In these

cases you must explain clearly to these sources:

- That you are undertaking research for the World Wide Web Foundation Open Data Barometer, a multi-country study of open data policy and practice due to be published in 2013;
- That any responses they give may be placed in a public open dataset, and that their responses will be included in this dataset;
- That they are under no obligation to respond to your questions, and can withdrawn from the interview at any time;

You should keep a clear written record, ideally with e-mail trail where interviews or consultations were arranged by e-mail, to show that you have asked each source to consent to the use of their responses in the Open Data Barometer.

1.3. Using the research platform

The research data collection is carried out using Google Spreadsheets. You must be online at all times during entering data, and we recommend using a modern browser such as Google Chrome, Firefox or the latest version of Internet Explorer.

To maintain anonymity of researcher and reviewers, we will provide you with a log-in account for accessing the survey documents.

Researchers

Getting started

You will receive an e-mail from the Open Data Barometer team providing you with a link to the survey spreadsheet and a username and password to access it.

This is a Google Docs username and password. If you already have a Google account then you will need to either:

- Sign-out and then sign-in with the log-in details we have provided;
- Open a new browser window in [Incognito mode](#) (Google Chrome only) which allows you to be logged in with different accounts in different browser windows;
- Use the 'Choose a different account' link, and then 'Sign in with a different account' link to be logged in with both your main and Open Data Barometer accounts at the same time.



You need permission to access this item.

You are signed in as **researcher2@opendatabarometer.org**, but you don't have permission to access this item. You can request access from the owner or choose a different account. [Learn more](#)

[Request access](#)

[Choose a different account](#)

DO NOT press the 'request access' button. You must use on the options above.

If you still cannot access the survey form after this it may be because your form have moved to another stage of the research process. For example, when it is in the review stage you will not have access to it.

You should store a copy of the link to the survey form safely.

You must use the online version of this spreadsheet. Do no download a copy to your computer or try to complete the sheet offline. You will need Internet access throughout your work on this spreadsheet.

	A	B	C	D	E	F	G	H	I	J
1	Open Data Barometer 2013 - Survey Form									
2										
3	Country Afghanistan									
4	Lead researcher Tim									
5	Reviewer Jose									
6	Completed? No									
7										
8										
9	This survey consists of 17 sections. Two with general questions, and 15 detailed dataset assessments. You can access each section from the tabs along the bottom of the screen. This page will be automatically updated as you fill in sections of the study.									
10										
11	Question Type	Question Group	Complete?	Ready for review?	Reviewed?	Accepted?				
12	Admin	Researcher Information Sheet								
13	General	Context questions	0 of 11	No	No	-				
14	General	Impact questions	0 of 6	No	No	-				
15	Dataset	Map Data	0 of 10	No	No	-				
16	Dataset	Land ownership data	0 of 10	No	No	-				
17	Dataset	Government service directory	0 of 10	No	No	-				
18	Dataset	Detailed census data	0 of 10	No	No	-				
19	Dataset	Detailed government budget	0 of 10	No	No	-				
20	Dataset	Detailed data on government spend	0 of 10	No	No	-				
21	Dataset	Company register	0 of 10	No	No	-				
22	Dataset	Legislation	0 of 10	No	No	-				
23	Dataset	Public transport timetables	0 of 10	No	No	-				
24	Dataset	International trade data	0 of 10	No	No	-				
25	Dataset	Health sector performance	0 of 10	No	No	-				
26	Dataset	Primary or secondary education performance data	0 of 10	No	No	-				
27	Dataset	Crime statistics	0 of 10	No	No	-				
28	Dataset	National environment statistics	0 of 10	No	No	-				
29	Dataset	National election results	0 of 10	No	No	-				
30										
31										
32										
33										
34										
35										
36										

+ ≡
Dashboard ▾
Researcher Information ▾
Context ▾
Impact ▾
D1 - Map ▾
D2 - Land ▾
D3 - Services ▾
D4 - Census ▾
D5 - Budget ▾
D6 - Spend ▾
D7 - Company ▾
D8 - Legislation ▾

When you first open the Barometer worksheet you will see a display as above. The dashboard sheet allows you to keep track of the question you have completed. Don't be daunted by the number to complete: the dataset questions should not take too long each.

DO NOT under any circumstances add rows or columns to the spreadsheet. You should only enter information in the cells marked for input.

You can move between the different data entry screens using the tabs along the bottom of the screen. If you cannot see all the sheets listed, the  button provides a list of them.

Entering data

Move through the question sections. The Context and Impact tabs ask 'To what extent questions'. Tabs D1 - D15 to contain questions on the availability of key datasets.

For each of the questions you will see a drop-down box for your score (highlighted in colour), and a box for '**Justifications and sources**'.

C1) To what extent is there a well-resourced open government data initiative in this country?

10	Scoring guidance	5	0
There is a strong national open data initiative with significant resources behind it, including dedicated staff and budgets. There is explicit commitment to open data from a senior government figure (e.g. Cabinet minister) and/or parliamentary backing for an open data initiative.		There is a small-scale open data initiative, or an open data initiative has been announced but is not yet resourced. Senior leadership is making commitments to increased government transparency, and/or some commitments to open data are being expressed by a junior minister / single ministry.	There is no evidence of a formal open data initiative, nor any commitment from government to release open data.

0

3

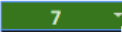
5

7

10

Justification and sources (include URLs)

You should reference and include links to policy statements on open data in addition to other source notes.

To provide a score, click the small downward arrow next in the coloured box. . When you have provided a score the box will turn green, and the county of question to complete on this page will be reduced.

Enter your justification for the score given, and details of the sources you can drawn upon in the Justification and sources box.

Note: Next to each question you will find a link to **Scoring Guidance** or '**Help**'. This will link to the relevant part of this Open Data Barometer handbook. **You should always read the**

guidance for the question you are answering in detail before you complete your answer.

Question	Yes/No?	Supporting information
a) Does the data exist?	-	- if YES: Please briefly describe the data (e.g. name it is known by; details about its quality and coverage) if YES: Which agency is responsible for the data? if NO: Please explain why this data is not collected or available at a national level in the country Justification and research sources
b) It is available online from government in any form?	-	- if YES: Where is the data found (URL)?

On the dataset assessment question sheets, you can provide a 'Yes/No' answer to each question from the drop-down box in the middle column .

These sheets also provide space to enter more detailed source information. Enter this into the cell under each heading. Do not type over the headings.

If you enter a long answer to some questions not all the text will display on the screen. This does not matter. You can check all the text is recorded by double clicking the cell and seeing that it is all there.

A reminder: do not add any new rows or columns to the sheets. Doing so may lead to your responses being rejected.

Flagging sheets for review

When you have completed a sheet of questions you can mark it as ready for review using the drop-down box in the information area at the top of the page.

Lead researcher	Reviewed by: (to be completed by reviewer only)	
Research carried out on:	Reviewed on (date):	
Dataset	Passed review?	
Map data	-	
Ready for review?		
No		
No		
Yes		
Question	Yes/No?	Supporting information
a) Does the data exist?	-	- if YES: Please briefly describe the data (e.g. name it is known by; details about its quality and coverage)

When you have marked all pages as ready for review **please e-mail your regional coordinator** so that they can start the review process.

Regional coordinators

Your regional coordinator will check your answers to ensure all areas are complete.

They may use the 'Comments' feature of Google Docs to highlight areas where you need to provide more details. See the information below on 'Reviewing and responding to comments'.

Once they are satisfied with your response they will forward it to a reviewer.

Reviewers

Reviewers will be invited to view the spreadsheet Google Doc as described in the 'Getting started' section above.

Reviewers **should not** edit any values on the sheets, with the exception of entering details of their name and the date of review on each sheet they have completed their review of this sheet.

Reviewers will:

- Check each of the sources given and ensure that this justifies the score given;
- Raise questions about any scores or answers which they feel may be mistaken, or open to alternative interpretations;
- Request further clarification if justifications do not provide enough detail, or if any details are missing.

To do this, reviewers will add comments to the document by:

- Selecting either the **Score** or the **Justifications and sources** cell, and choosing 'Insert > Comment' from the menus; 
- Typing a comment highlighting the required changes, or questions for the researcher.

Researchers may receive these comments by mail. You can change whether or not you receive comments by mail by changing your 'Notification' settings from the comments button at the top-right of the screen.



If the reviewer is happy that all the scores and justifications on a sheet are provided to a high

standard then they should select 'Yes' from the 'Passed review' drop-down box at the top of the sheet.

When the reviewer has completed a review of all the sheets they should **notify the regional coordinator by e-mail**.

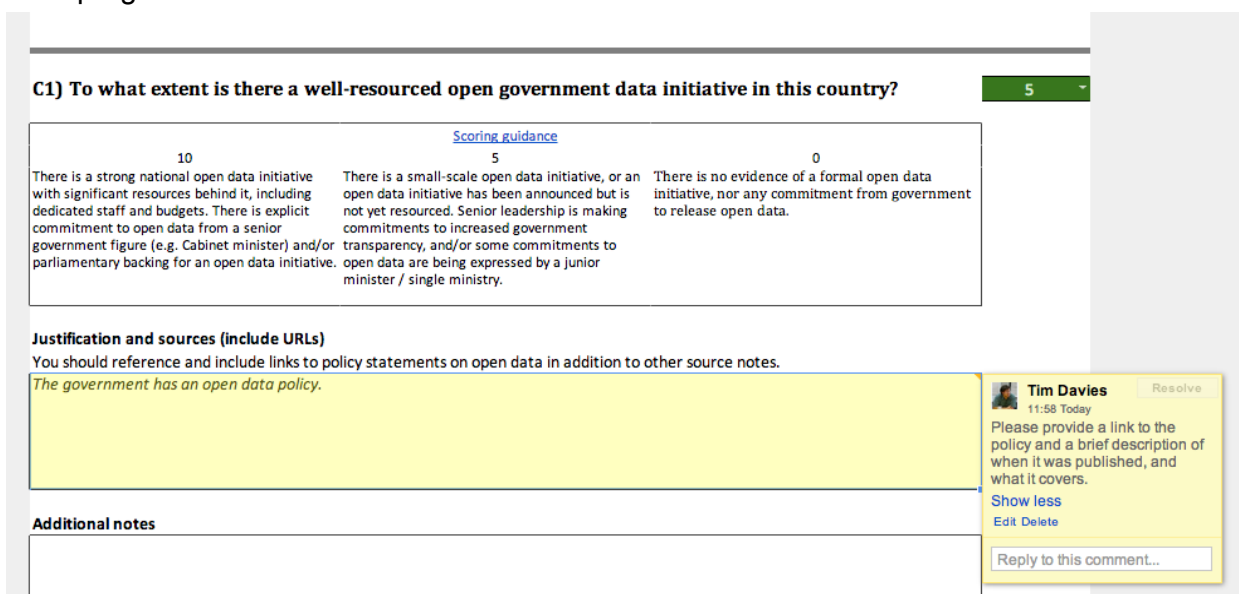
Researchers

Reviewing and responding to comments

After each review, the researcher will be sent an e-mail to invite them to return to the sheet and address all comments.

Cells with comments on are indicated by a small orange triangle in the top corner . A list of all the comments is also provided from the 'Comments' button at the top-right of the screen, and a number on each tab indicates how many comments there are on that sheet.

To view a comment, click on the cell that comment is against, or click in the comments listing at the top-right of the screen.



C1) To what extent is there a well-resourced open government data initiative in this country? 5

10	5	0
There is a strong national open data initiative with significant resources behind it, including dedicated staff and budgets. There is explicit commitment to open data from a senior government figure (e.g. Cabinet minister) and/or parliamentary backing for an open data initiative.	There is a small-scale open data initiative, or an open data initiative has been announced but is not yet resourced. Senior leadership is making commitments to increased government transparency, and/or some commitments to open data are being expressed by a junior minister / single ministry.	There is no evidence of a formal open data initiative, nor any commitment from government to release open data.

[Scoring guidance](#)

Justification and sources (include URLs)
You should reference and include links to policy statements on open data in addition to other source notes.

The government has an open data policy.

Additional notes

Tim Davies 11:58 Today Resolve

Please provide a link to the policy and a brief description of when it was published, and what it covers.

[Show less](#)
[Edit](#) [Delete](#)

Reply to this comment...

You should read the reviewers comment carefully, and identify any changes required to your score, or to the justification and source information you have provided.

Either:

- **If you agree with the comment:** make the required changes, and then write a brief response in the 'Reply to this comment' box, and press 'Reply';
- or**
- **If you do not agree with the comment:** use the 'Reply to this comment' box to explain

your opinion, and press reply to send your response to the reviewer;

Do not press the 'Resolve' button.

When you have addressed all the comments, **e-mail the regional coordinator** to let them know that your sheet is ready for a final review.

Final review

The regional coordinator will check your responses, reviewing each comment and response.

Where comments have been satisfactorily resolved they will click the 'resolve' button against the comment.

Where comments have not been resolved, they may either add extra notes directly, or may forward the sheet back to the reviewer to carry out a further review.

The reviewer may resolve comments, or add additional responses if there are still areas of disagreement.

Support

If you have any questions about the research process at any point you should contact your regional coordinator, or tim.davies@webfoundation.org.

1.4. Justifications and sources

Each question provides space for '**Justification and sources (including URLs)**'. Please enter these details **only** in the box provided. **Do not** add any extra lines to the spreadsheet. If you are entering multiple sources separate responses with a full stop and/or space.

You **must** provide your justification and sources, and notes in **english**. The resources linked to can be in other languages, and you should research and search for sources in the official languages of the country you are researching wherever appropriate.

You should **always** provide a justification, even when scoring a question zero (0), or answering No (unless in the dataset assessment you have also answered No to the preview two questions). In these cases you must explain how you tried to locate the details requested in the question. For example, describe the searches you tried to find an open data policy or a particular dataset which ended with no evidence of a policy or that dataset being available.

An edited copy of the source information provided may be published in the open dataset alongside the data of the Open Data Barometer.

Space is also provided for any **additional notes**. Additional notes will be available to the

reviewer, but will not be published. You can use this space to note if you have any areas of uncertainty about a question response, or to record notes to yourself for following up this question further.

You may draw upon the following kinds of sources:

- **Laws and regulations:**

- Researchers should identify the law (full name, article, etc.) and provide a direct quote or a detailed explanation of the law in the Comments box, as needed.
- Researchers must refer, where appropriate, to case law, specific legal articles or statutes.
- Researchers are to reflect and describe applicable traditional and customary law when necessary.
- Laws cited must be domestic laws rather than international law.
- Please use this format when citing laws or regulations: Name of Law, Article, Section Number, Year, Hyperlink to law. Example: *Constitution Acts, Part I: Canadian Charter of Rights and Freedoms, Section 6. 1982.*
http://laws.justice.gc.ca/eng/Const/page-11.html#sc:7:s_6

- **Web sites and searches**

- Cite websites using the full URL, starting with http(s)://, and include the date the website was accessed.
- If the information can only be accessed behind a search interface, please provide details of how a reviewer could access the same pages that you use to justify your answer.

- **Media articles**

- Articles should be from the reporting period of June 2012 to June 2013, or referring to events during this timeframe.
- Please use the following format when citing media articles: name of author, name of publication, title of article, date published, and a URL/hyperlink.
- If the article is not available online, please save a PDF or scanned copy and note this in the additional notes field. You will be given details of how to upload this scanned copy during the review process.

- **Interviews, e-mail conversations, web content, mailing list content and other desk research**

- Score choices should reflect the project's official reporting period of June 2012 to June 2013. Sources and score choices that fall outside the official study period by a few weeks are also allowed; beyond that, however, sources and score choices that fall outside the study period should not be included.
- Exercise professional judgment in determining whether opinions of an interviewee

are factual and accurate. We strongly suggest researchers corroborate information obtained in interviews with desk research.

- Please use the following format when citing interviews: Interview sources must include the full name of the interviewee, the name of the interviewee's employer, job title, date and place or method (city, country / e-mail or skype) where the interview was conducted. Example: *Jane Doe, Ministry of Justice, Director General, June 22, 2012, Kampala, Uganda.*
- Anonymity: If it is not feasible to publish the name of an interviewee (out of justified fear for the interviewee's safety or negative professional ramifications), please include the name of the interviewee in the "Additional notes" box in Indaba. In the "Justification and sources" box, simply state "Anonymous" for the interviewee name while providing as many other details as possible (e.g. Anonymous, Ministry of Finance, government official, June 1, 2012, Cape Town, South Africa).

- **Journal Articles**

- Please use the following format when citing journal articles, written documents or other third-party desk research: provide the full citation and whenever possible please provide a URL/hyperlink (or attach an electronic copy if a link is unavailable). Example: *Comstock, E. W. & Butler, J. W. (2000) Access denied: The FCC's Failure to Implement Open access as Required by the Communication Act. Journal of Communications Law and Policy, Winter.*

Researchers' comments are expected to offer a brief and reasoned argument to explain the score choice. Make reference to concrete examples or details obtained from the research (interviews, reports, etc.). General or vague comments are not acceptable and will be returned to the researcher for improvement.

An isolated example is not enough evidence with which to select a score. Researchers need to consider and assess the overall context before selecting the appropriate score.

Personal opinions and broad generalizations without specific supporting evidence within the period of study are not acceptable.

1.5. Confidence level

For each of the Context and Impact questions, and each overall dataset question, you can provide a 'Confidence Level' of 0%, 25%, 50%, 75%, 95% or 100% for how certain you are of the answer you have given.

You should aim for an average confidence level of above 75% in each section.

1.6. Additional research guidance

Getting acquainted with the question formats

You should attempt one of the dataset assessment questions early in your research process to check you are happy with your ability to use the 10-point question scale these use. You can ask your project manager or the research coordinator to review your first attempt at one of these questions if you are unsure of the answers you have given.

Access to literature

If the sources you need are in academic journals which you do not have access to, due to paywalls or other factors, please contact your project manager with full details of the article and we will explore options to obtain it for you.

General definitions

Defining Open Data

Open data involves “(1) the proactive disclosure of information; (2) the internet being the primary medium for such disclosure; (3) information being made available for access and for re-use free of charge and; (4) information being made available in a machine-readable format to enable computer-based re-use.”¹

The [Open Knowledge Definition](#) is often used to provide a formal definition of what it means to have open data, stating that “A piece of data or content is open if anyone is free to use, reuse and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike.”

A number of other definitions or lists of principles have been put forward that set out criteria for good quality open (government) data, such as a [Sebastapol Principles](#).

The 10-point checklist used for dataset assessment questions picks out key features that these different definitions and principles use to distinguish the openness of a dataset, making it possible to both judge datasets on a binary open/closed basis, and to judge degrees of openness.

What open data is not:

There are many other policy areas that can overlap with open data. However, it is important to understand how they are distinct from open data. For example, open data is conceptually distinct from:

- **Open government** - whilst an open government policy might mention open data, the two are not identical. Always check for explicit discussion of open data;
- **E-government** - policies to place government services online *might* have an open data element to them, but in many cases they only give citizens access to specific services or small extracts of data, rather than providing full access to machine-readable data for citizens to access.
- **Data sharing** - governments may increase data sharing between departments, but with limits on who the data is shared with, or who can re-use it. It is only open data when anyone can re-use it;
- **Open Access** - open access focusses on access to (academic) documents and publications rather than datasets;

¹ Open Government Data (Centre for Internet and Society, India) in Transparency & Accountability Initiative ‘[Opening up Government](#)’

² The Open Definition: <http://opendefinition.org/okd/>

For further introductions to open data see the following videos:

- **Open Government Data and the OD4D project (Spanish):**
<http://www.youtube.com/watch?v=lzg9qmP9Nxl>
- **What is Open Data** <http://www.youtube.com/watch?v=BkdRQipgodE>

Section 2: Question guidance

Context questions

You should spend an absolute maximum of one hour on each question (most should take 15 - 30 minutes of desk research, plus additional time consulting with expert informants). If after in-depth searching you have not found evidence to support your scores then you can use the confidence level to indicate how certain you are of the score you are giving.

Remember: scores of **zero** also need 'Sources and justification' explaining how you are attempted to locate evidence on the relevant area of context, and detailing any informants you have consulted for this question.

C1: To what extent is there a well-resourced open government data initiative in this country?

SCORING CRITERIA

10	5	0
There is a strong national open data initiative with significant resources behind it, including dedicated staff and budgets. There is explicit commitment to open data from a senior government figure (e.g. Cabinet minister) and/or parliamentary backing for an open data initiative.	There is a small-scale open data initiative, or an open data initiative has been announced but is not yet resourced. Senior leadership is making commitments to increased government transparency, and/or some commitments to open data are being expressed by a junior minister / single ministry.	There is no evidence of a formal open data initiative, nor any commitment from government to release open data.

SCORING GUIDANCE

Look for evidence of **an explicit government commitment to open data**.

An open data initiative is a plan by the government to release government data online to the public. It has four main features:

1. The government discloses data or information without request from citizens. This may be according to a release schedule or *ad hoc*.
2. The Internet is the primary means of disclosure. Mobile phone applications may also be used for disclosure.
3. Data is free to access and re-use, e.g. open licenses;
4. Data is in a machine-readable format to enable computer-based reuse, e.g. spreadsheet formats, Application Programming Interface (API).

Look for all these features in the policy you are assessing for it to receive the maximum scores.

The ten principles of open government data are:

1. Completeness
2. Primacy
3. Timeliness
4. Ease of Physical and Electronic Access
5. Machine readability
6. Non-discrimination
7. Use of Commonly Owned Standards
8. Licensing
9. Permanence
10. Usage Costs

For more information on what these principles entail see:

<http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

Significant resources for an open government data initiative include a sufficient budget, personnel and facilities to carry out the mandate of the open data initiative, including technical personnel with appropriate qualifications for dealing with open data issues.

Note that this question is only concerned with initiatives led by the **national government**. Open data initiatives covering the country, but organised by a third party, such as the African Development Bank or another regional organisation should not be counted, although you can detail these in the 'Additional notes' section.

SOURCE GUIDANCE

- The existence of a government open data action plans, policy or directive
- Speeches by government leaders about open data
- Conversations with members of the open data movement, both government and civil society
- Conversations with the “civic hacker” community
- Open Government Partnership country action plans (www.opengovpartnership.org) that

contain explicit commitments to open data

- For countries in Latin America, the OD4D project has compiled a database of policies and practice at <http://www.od4d.org/2012/11/23/mapping/>. Ensure you check for updates more recent than this database reports.
- Reports published by the media, academic and policy journals, and development and multilateral bodies (e.g., World Bank Open Data Readiness Assessment, IFC, OECD, African Development Bank).

C2: To what extent does the country have a functioning right-to-information law?

SCORING CRITERIA

10	5	0
Citizens generally receive responses to requests for government information within a reasonable time period and at a reasonable cost, and responses are typically of acceptable quality. An RTI law or similar legal right enshrines the right to such requests.	Citizens may not receive timely responses to requests for government information, or responses are missing key information. The right to request government information may be generally guaranteed by a vague constitutional right but is not further protected by a dedicated law or enforceable regulation.	Requests for government information are generally not responded to at all, or responses are of extremely low quality. There may be no right-to-information law at all.

SCORING GUIDANCE

This indicator addresses whether the disclosure requirements are “effective.” The basic requirements for them to be considered “effective” are whether information is:

1. available to the public for free or at reasonable/minimal costs in a variety of venues (e.g., online, government agency offices);
2. can be accessed by citizens within 30 days, and
3. answers the specific request, with explanations for refusal to release information;

For a 10 score, there can be exceptions in which information is not released because of national security or individual privacy concerns (e.g., medical records, sexual orientation, etc.), but the legal reason must be stated clearly in the response from the government to the citizen who requested the information.

SOURCE GUIDANCE

- A non-exhaustive databases of right-to-information laws can be found at RTI Rating [www.rti-rating.org] and from the Public Accountability Mechanisms [www.agidata.org/pam] database
- You can check existing assessments of RTI laws in practice: Global Integrity, Category I-3 [www.globalintegrity.org]
- Conversations with NGO officers with expertise in access-to-information issues and investigative journalists
- Reports published by the information agency, media reports and publications by development/donor agencies

C3: To what extent is there a robust legal or regulatory framework for protection of personal data in your country?

SCORING CRITERIA

10	5	0
A legal or regulatory data protection framework exists that is broadly applicable, provides the right of choice/consent to individuals, provides the right to access and/or correct one's personal data, imposes clear responsibilities on information holders and provides a right of redress against both private and public bodies that violate data privacy.	A legal or regulatory regime exists but is missing some of the key elements understood to promote best practices around data protection policies, including broad applicability, the right of choice/consent to individuals, the right to access and/or correct one's personal data, clear responsibilities on information holders, and/or the right of redress against both private and public bodies that violate data privacy.	A legal or regulatory regime to promote data protection does not exist or is so devoid of precision and/or the understood best practices as to render it useless in practice.

SCORING GUIDANCE

Strong data protection regimes include the key features noted below, namely broad applicability, the right of consent, the right to access and correct the information, obligations on data controllers, and the right of redress.

1. **Broad applicability** – these rules should apply to personal data sets and data controllers in both the public and private sectors.
2. **The right of choice/consent** – Individuals should normally be given the choice of whether their information is collected. There should be only limited exceptions to this where there is an overriding interest, defined in law, in the collection of such information. This implies that individuals understand and are given clear notice of a public or private body's information practices before any personal information is collected. This notification should describe what information is proposed to be collected and held, who will collect it, how the information will be used and who will have access to it. It should also be clear to the subject whether the provision of the requested information is voluntary or required by law and of the consequences of refusing to provide the requested information. Information should not be used for purposes that are incompatible with the use for which the information was originally collected.
3. **The right to access and correct** – Individuals should have the right of access to any

information held about them at reasonable intervals and without undue delay. They should also have the right to require the data controller to correct any inaccuracies or to delete the data, where appropriate.

4. **The responsibilities of information holders** – Data controllers must take reasonable steps to ensure that the information they hold is accurate and secure. Access to the data should be limited in accordance with the established uses of the data. Transfers should be made only to third parties that can ensure similar respect for data protection principles. Data should be destroyed once the information is no longer needed for the established uses, or converted to anonymous form. While information is held, appropriate steps should be taken to ensure the confidentiality, integrity and quality of the data.
5. **The right of redress** – Individuals should have the right of redress against public and private bodies that fail to respect data protection rules in relation to data about them. Remedies can be provided through self-regulation, private law actions and government enforcement. Oversight of the system should be undertaken by an independent body.

SOURCE GUIDANCE

- This 2012 paper (Global Data Privacy Laws: 89 Countries, and Accelerating) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2000034 provides a list of 89 countries with data protection laws and the names of the laws to assist in researching whether a law is effectively implemented.
- Privacy International maintain country reports at <https://www.privacyinternational.org/reports/country> (note, these are not always very recent)
- Conversations with experts in data privacy, information privacy or access to information, such as academics, researchers, think tanks and NGOs.
- Reports published by the media, academic and policy journals, and development and multilateral bodies (e.g., World Bank, IFC, OECD, African Development Bank).

C4: To what extent are civil society and information technology professionals engaging with the government regarding open data?

SCORING CRITERIA

10	5	0
There are coordinated campaign(s) calling for open data or working with governments to promote open data. There are several engaged groups using open government data, and outreach events and activities may be conducted to promote use of open data. The government regularly engages with the user community.	There is some demand or engagement over open data from individuals or communities, but it may be isolated. There is no coordinated campaign for open data usage, and the government does not engage the community regularly.	There is little, if any, evidence of demand or use of open data.

SCORING GUIDANCE

Campaigns for open data are often composed of civil society organizations, data technologists, informational professionals, computer experts and ordinary citizens who advocate for greater access to government data. A 10 score indicates that government officials recognize these organized campaigns and engage in discussion with community leaders about which data to release, when and in what forms.

SOURCE GUIDANCE

- Online evidence of an open data community, including reports of events and other activities;
- Conversations with open data specialists in civil society organizations or individuals who are directing open data campaigns;
- Conversations with government officials working in open data offices or projects;
- Conversations with NGO officers with expertise in open-data and access-to-information issues, and investigative journalists;
- Reports published by the media, academic and policy journals, and development and multilateral bodies (e.g., Open Knowledge Foundation, Sunlight Foundation, etc.);

C5: To what extent are city or regional governments running their own open data initiatives?

SCORING CRITERIA

10	5	0
A majority of the largest cities and regions in the country have open data initiatives. A significant number of these have strong political backing and are well resourced.	A number of different cities or regions across the country have open data initiatives. These may be predominantly run by civil society, or may have limited resources dedicated to them. Open data initiatives exist in cities and regions beyond the national or state capitals.	There are no city or regional open data initiatives.

SCORING GUIDANCE

Open government data does not just involve central government. Regional, city and local government may all adopt open data initiatives.

The criteria of an open data initiative are the same as in C1.

SOURCE GUIDANCE

- Conversations with open data specialists in civil society organizations or individuals who are directing open data campaigns;
- Conversations with government officials working in open data offices or projects;
- Conversations with NGO officers with expertise in open-data and access-to-information issues, and investigative journalists;
- Searching for data portals and platforms run by cities or regions;
- Reports published by the media, academic and policy journals, and development and multilateral bodies (e.g., Open Knowledge Foundation, Sunlight Foundation, etc.);

C6: To what extent is government directly supporting a culture of innovation with open data through competitions, grants or other support?

SCORING CRITERIA

10	5	0
The government has made a substantial commitment to support a culture of innovation with open data, including significant financial incentives and supporting a range of different activities. These include events targeting private sector use of open data. Multiple government departments are involved in supporting innovation with open data.	There are a range of interventions to support a culture of innovation. For example, there are three or more examples of competitions, funding schemes or hackathon events run by government AND two or more departments or agencies are involved in running these schemes.	There is no government support for innovation using open data.

SCORING GUIDANCE

Governments can adopt a range of approaches to stimulate a culture of innovation around open data including:

- **Running competitions** in which prize money is offered to innovators creating tools, services or commercial applications using open data;
- **Organising [hackathon](#) events** which invite developers to create prototype tools and services over one or two day events;
- **Organising incubators and open data boot camps** specifically targeted at supporting innovative uses of open data;
- **Offering grant funding or [innovation vouchers](#)** specifically targeted at encouraging businesses to engage with open data;

Sometimes these are run as one-off activities, organised with minimal budgets by small groups of staff. In these cases, a maximum score of 5 should be given. To receive a score of 10 there should be clear evidence of government dedicating investment to support innovation with open data, including support to private sector re-users of open data.

If you find a number of non-government activities to support innovation with open data you may include details of these in the additional notes field, but only events with significant government involvement should feature in the score you give.

SOURCE GUIDANCE

- Conversations with open data activists, government officials involved in open data, or entrepreneurs working on open data in the country;

- Search for government press releases and announcements about support for open data innovation activities;
- Search for donor press releases, project funding documents and announcements of support for open data innovation activities;
- Search for 'hack days', 'open data boot camps', 'open data competitions' and check to identify if any of these were supported by government;

C7: To what extent is training available for individuals or businesses wishing to increase their skills or build businesses to use open data?

SCORING CRITERIA

10	5	0
There is widespread access to high quality training covering the full range of open data issues, such as data technology, data science and statistics, data visualisation and the business and legal aspects of open data. Both basic and advanced training are available from a range of different providers and in a range of cities and locations across the country. Much of the training provided is 'world class' and would equip those trained to operate in a global marketplace.	There is limited access to training on open data issues, such as data technology, data science and statistics, data visualisation and the business and legal aspects of open data. Whilst there may be some examples of training that would equip those trained to operate in a global marketplace, the majority of training may be of middling or low quality.	There is very limited training available, or those wishing to learn have to access training from outside the country.

SCORING GUIDANCE

Working with open data involves a wide range of knowledge and skills, including:

- **Web technologies** - covering issues such as web development, publishing data on the web, and building applications with open data for a range of platforms, including mobile;
- **Data science** - covering core statistical skills, and the ability to work with large online datasets from heterogeneous sources;
- **Data visualisation** - covering design and programming skills to create interactive and static data visualisations and infographics;
- **Legal aspects of open data** - addressing issues such as the licensing of open datasets, and legal issues in open source, open data and open knowledge;
- **Business aspects of open data** - addressing the creation of commercial enterprises on top of open data or open source - delivered through formal training or business mentoring and incubation programmes;

Training may be delivered through i), both full-time and part-time, or through professional development courses. Training may also be delivered through business incubator programmes, or short-term boot-camp training events.

Where training events are one-off, or short-term donor funded interventions in one or two locations in the country, the maximum score available should be. Where training on the majority of the topics noted above is accessible to businesses and individuals who may be interested in it, then a score of 10 should be given.

Where online training is available in languages accessible to businesses and individuals in the country, *and* there is evidence that this training is used by nationals of the country, then this may count towards your assessment of 'limited access to training' (5), but online only courses should not be taken to count towards 'widespread access to training' (10).

Tip: You may wish to complete question C7 first, before returning to complete your assessment of C6.

SOURCE GUIDANCE

- Conversations with developers and technology businesses interested in developing their open data capacity to identify the training they access;
- Conversations with training providers about the courses offered and the range of courses available;
- Course directories and promotional material from further and higher education providers;
- Web searches for training providers on the particular topics noted above;

C8: To what extent are the following forms of training available in the country?

SCORING CRITERIA & GUIDANCE

For each question choose from the three options:

- **No Provision**
- **Limited Provision**
- **Widespread Provision**

For each answer where there is limited or widespread provision you should give one or more example web link pointing to details of a training course online, or a directory listing these training opportunities. Where no online materials to justify your score are available, please upload a supporting document to your working folder.

You should consider online courses which are actively used by nationals of the country as 'limited provision'. You should only give a score of 'Widespread Provision' where the majority of people with an interest in training in the subject have a reasonable chance of accessing training (subject to having the funds available to pay for it).

The cost of training should not factor in your assessments of the breadth of provision. However, if cost of training provision (or other factors, such as geographic distribution of opportunities) creates significant barriers to certain groups accessing training then you may detail this in additional notes.

The kinds of training you are asked to assess are:

- **Web technologies** - covering issues such as web development, publishing data on the web, and building applications with open data for a range of platforms, including mobile. The training descriptions do not need to explicitly mention open data;
- **Data science** - covering core statistical skills, and the ability to work with large online datasets from heterogeneous sources. This may talk about 'big data' or 'digital data' and does not have to explicitly mention open data;
- **Data visualisation** - covering design and programming skills to create interactive and static data visualisations and infographics. The training does not need to explicitly mention open data, but should focus on **data** visualisation, rather than just graphics.
- **Legal aspects of open data** - addressing issues such as the licensing of open datasets, and legal issues in open source, open data and open knowledge. The training description should explicitly mention open data;

- **Business aspects of open data** - addressing the creation of commercial enterprises on top of open data or open source - delivered through formal training or business mentoring and incubation programmes. The training descriptions should explicitly mention open data;

SOURCE GUIDANCE

- Conversations with developers and technology businesses interested in developing their open data capacity to identify the training they access;
- Conversations with training providers about the courses offered and the range of courses available;
- Search the course directories and promotional material from further and higher education providers;
- Web searches for training providers on the particular topics noted above;

C9: To what extent are academic institutions in the country opening up their data?

SCORING CRITERIA

10	5	0
There is a national policy requiring that datasets generated through state-funded academic research are published as open data wherever possible. There are well resourced national data archives to support publication of open data from research. Other datasets from across academia such as course directories and performance statistics are also published as open data.	There is a national policy on archiving research data, but it does not specifically require open licenses. A number of Universities provide data archives, but there are not necessarily national archives in place. A small number of Universities are publishing additional open datasets, but the practice is not widespread.	There is very limited publication of open data from academia.

SCORING GUIDANCE

Open data is not just an issue for central governments. Open data in science and academic research is also an important topic.

A number of governments have introduced policies requiring state funded research to be deposited with 'data archives' for long-term preservation, and Universities or national disciplinary programmes are increasingly developing their own thematic data archives. This question explores whether, in addition to 'open access' policies (which generally focus on access to publications) or 'data archival' policies, academic institutions are also implementing open data policies that lead to the publication of datasets under open licenses.

This question also asks about whether academic institutions are publishing administrative data as open data. For example, www.data.ac.uk in the United Kingdom brings together open data [published by a small number](#) of UK universities on facilities and equipment.

Institutions may also be publishing Linked Open Data.

This question is focussed on an emerging issue, and so you should not spend more than 30 minutes researching for this question. It is unlikely that many countries will score above 5 on this question in 2013.

SOURCE GUIDANCE

- Check the 'Registry of Research Data Repositories' <http://www.re3data.org/> for any repositories in the country;
- Search for Open Data Repositories in the academic domain of the country (e.g. UK Universities all have websites under .ac.uk domains, so a Google search of 'open data site:ac.uk' highlights examples of open data);
- Have conversations with open science, open access and open data activists and campaigners in the higher education sector: make sure you focus on **open data** rather than more general **open access** issues;
- Check for government or research funding body policies on 'open access', 'research data management': look specifically for mention of open licenses;
- Consult with or check the archives of open science mailing lists and communities;

C10: To what extent are civil society in the country opening up their own data?

SCORING CRITERIA

10	5	0
There is widespread publication of open data by civil society from a number of different sectors. There are examples of open data publication from both larger and smaller civil society organisations. Publishing open data is seen as a default behaviour for civil society. The majority of data is provided under clear open data licenses.	There are a number of examples from different sectors of civil society publishing open data, including details of funding flows to civil society and data from research carried out by civil society. At least one large civil society organisation is publishing multiple open datasets. A number of these examples may not have perfect open data licenses, but it is clear that the data was intended for reuse	No examples of civil society publishing open data could be located.

SCORING GUIDANCE

Civil society organisations hold a range of datasets that can be made available as open data. Projects like the [International Aid Transparency Initiative](#) or [hGrant](#) are encouraging charities to publish details on their funding, or encouraging foundations to publish open data on the grants they give. Charities can also publish data collected through research as open data.

For the purpose of this question you may consider journalism to be a sector of civil society, such that if news outlets are publishing open data from their own research this would count as an example of one civil society sector publishing open data.

To qualify as open data the datasets a charity publishes should be:

- Provided online;
- Provided in machine readable formats;
- Providing under an open license;

At present there are relatively few examples of charities and civil society organisations publishing open data.

The terms 'larger' and 'smaller' civil society organisations are recognised as being relative to the country. I.e. whilst in some countries a large CSO would have hundreds of employees, in other countries the larger CSOs are 25 - 30 people. Judge based on the relative sizes of CSOs in the country itself - rather than based on global comparisons of CSO scale.

Do not include examples where charities or civil society organisations are simply republishing or visualising government datasets. However, you can include cases where civil society organisations or networks (including informal networks) have re-publishing *as open data* information that was originally obtained from government in non-data formats (for example, a paper response to a Freedom of Information request that has been re-typed into a database form)

This question is focussed on an emerging issue, and so you should not spend more than 30 minutes researching for this question. It is unlikely that many countries will score above 5 on this question in 2013.

SOURCE GUIDANCE

- Conversations with open data activists, advocates and service providers;
- Checking open data catalogues such as The Data Hub (<http://datahub.io/>), or other non-government open data catalogues for datasets from civil society organisations: check to see how many different sectors the data is drawn from, and whether this is a one-off publication or data, or whether publication appears to be sustained, and part of 'business as usual'
- SlashOpen - <http://slashopen.net/directory/>
- Web searches, blog posts, press releases and news articles discussing charities or civil society releasing open data - you should check to verify that the dataset discussed really is open data;

C11: To what extent are businesses in the country opening up their own data?

SCORING CRITERIA

10	5	0
There is widespread publication of open data by companies from a number of different industries. There are examples of open data publication from market leaders, small/medium enterprises and micro businesses. The majority of examples are from sectors other than the ICT sector.	There are examples from multiple different industries of companies opening up their own data. These examples include at least one market leading firm, as well as examples from small/medium enterprises and micro-businesses. Many of the examples available may be from the ICT sector.	No examples can be found of companies in the country publishing data as open data.

SCORING GUIDANCE

Open data is an idea that can apply to companies as well as to governments. Businesses can publish a wide range of datasets as open data, including:

- **Product catalogues and inventories**
- **Timetables and service details**
- **Environmental performance data**
- **Research data**
- **Geographic data**

To qualify as open data publication a business should be:

- Providing the data online;
- Providing it in machine readable formats;
- Providing it under an open license;

Where not license is given, or the license for access to the data restricts how a third-party can reuse it then the data **does not** count as open data.

At present there are relatively few examples of businesses publishing open data. For example, whilst large Internet firms like Twitter or Foursquare might provide machine readable access to their data, this is not open data because the licenses they provide the data are not open licenses: they contain terms that restrict how the data may be re-used.

This question is focussed on an emerging issue, and so you should not spend more than 30 minutes researching for this question. It is likely that very few countries will score

about 5 on this question in 2013.

SOURCE GUIDANCE

- Conversations with businesses that are active around open data;
- Web searches, blog posts and press releases discussing a business releasing open data - you should check to verify that the dataset discussed really is open data;
- Checking open data catalogues for datasets from businesses;

Impact questions

I1 to I5 ask you to assess the degree to which there is any evidence that **open data release by the country government** has had impacts in a variety of different domains in the country.

The question scoring criteria invite you to look for credible claims made in academic publications, mainstream media and online sources which explicitly attribute certain impacts to open data released by the country's government. The highest scores are only available where there is peer-reviewed or audited evidence of impact.

We anticipate that in the 2013 Barometer there will be very few countries that score the highest scores, and some countries may have very low scores on this question.

Question I6 asks about entrepreneurial use of open data.

A note on 'impact'

Measuring impact is notoriously difficult. Establishing a solid causal connection between open data and particular changes is clearly beyond the scope of a survey such as the Open Data Barometer. However, for the purpose of the Barometer, claims made in credible sources concerning possible impacts of open data are a useful proxy indicator of areas in which impact may be occurring, and to allow initial comparison between countries.

Claims about impact go beyond simple descriptions of where open data has been used. For example, a newspaper might report that open data has been used to create a bus timetables application, to visualise the budget of the country, or in a hackathon focussed on public services. It is only when they relate this to some other outcome, such as budget savings, or an increase in use of public transport, that an impact claim has been made.

In future iterations of the Open Data Barometer we will be looking at additional methods for tracking potential impacts of open data, building on the evidence gathered through these questions in 2013.

With the exception of question I6, for each topic area you should carefully check to ensure the impacts you cite can be reasonably attributed to **open government data from the country in question**. For example, a crowdsourcing application created by an NGO to monitor public services may contribute to more efficient government, but if it does not use open data from government then, for the purposes of these questions, you should not include it in your analysis. By contrast, if a government publishes data on its spending, and there are cases of third-parties using this to highlight where government could use its resources more efficiently, this would count as a case where impact could be cited from open government data.

Where examples draw on open data from outside the country (e.g. an ICT based in the country is using open data from the World Bank, some other government, or some global dataset) you may report this in Justifications and Sources, but these examples should not count towards scores higher than 5, as higher scores are only available where the data being used comes from the national government's open data releases.

The top score (10) should only be given where at least one credible peer reviewed article clearly attributes measurable impacts to open data *and* there are further cases of impact in other credible media sources and blog posts etc.

Remember to focus on cases of open data released by national government in making your score assessments. You may include details of city level examples as additional notes in your justification.

The justification of 'widespread impacts' should include examples cited from at least three different sectors (e.g. health, education, transport).

I1: To what extent has open data had a noticeable impact on increasing government efficiency and effectiveness?

SCORING CRITERIA

10	5	0
Open data is widely cited to have made a significant contribution to government efficiency, with rigorous evidence to back these claims: for example a peer-reviewed study or government audit showing the contribution of open data to government efficiency.	Two or more cases in the media or credible online sources where open data is cited to have contributed to increased government efficiency	No evident impact at all

SCORING GUIDANCE

Open data could lead to improvements in government efficiency and effectiveness in a number of ways:

- By enabling government departments to better plan and target resources;
- By allowing outside actors to scrutinise government use of resources and highlight areas for savings;
- By enabling outside actors to build new services on top of open data which deliver more effective public services;
- By supporting collaboration between different government departments;

You may find other ways that open data impacts government efficiency and effectiveness in addition to these.

SOURCE GUIDANCE

- Conversations with open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media
- Blog posts and forum discussions;

I2: To what extent has open data had a noticeable impact on increasing transparency and accountability in the country?

SCORING CRITERIA

10	5	0
Open data is widely cited to have made a significant contribution to increased transparency and accountability, with rigorous evidence to back these claims: for example, a peer-reviewed study showing a positive impact on transparency and accountability.	Two or more cases in the media or credible online sources where open data is cited to have contributed to greater transparency and accountability.	No evident impact at all

SCORING GUIDANCE

Open data could lead to improvements in government transparency and accountability in a number of ways:

- Through supporting journalism and data journalism which uncovers wasteful spending, corruption or other wrongdoing by government departments or officials;
- Supporting the creation of applications which allow citizens to report on their experience of government services (for example, when a directory of schools or hospitals helps third-parties build a school or healthcare performance reporting application for citizens);
- Supporting scrutiny of government decision making;
- Supporting greater citizen engagement in policy making;

You may find other ways that open data impacts government transparency and accountability in addition to these.

SOURCE GUIDANCE

- Conversations with open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media;
- Blog posts and forum discussions;

I3: To what extent has open data had a noticeable impact on environmental sustainability in the country?

SCORING CRITERIA

10	5	0
Open data is widely cited to have made a significant contribution to the environmental sustainability, with rigorous evidence to back these claims: for example a peer-reviewed study showing the impact of an open dataset on environmental sustainability.	Two or more cases in the media or credible online sources where open data is cited to have had an impact on environmental sustainability.	No evident impact at all

SCORING GUIDANCE

Open data could lead to impacts on environment sustainability in a number of ways:

- Through enabling greater scrutiny of pollution impacts or environmental impacts of government projects or private enterprise;
- Through supporting greater attention to be paid to environmental factors in planning projects;
- Through encouraging government buildings to make more efficient use of energy;
- Through raising citizens awareness of their own environmental impacts;
- Through supporting campaigns on environmental issues;

You may find other ways that open data impacts environmental sustainability in addition to these.

You may include in your assessment open datasets which, whilst not directly published by government, are published as a result of a government mandate. For example, where a regulator may requires that factories publish pollution statistics. However, you should check that the published data is **open data**, licensed for reuse in these case.

SOURCE GUIDANCE

- Conversations with open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media;
- Blog posts and forum discussions;

I4: To what extent has open data had a noticeable impact on increasing the inclusion of marginalised groups in policy making and accessing government services?

SCORING CRITERIA

10	5	0
Open data is widely cited to have made a significant contribution to the inclusion of marginalised groups, with rigorous evidence to back these claims: for example a peer-reviewed study showing the greater inclusion of at least one marginalised group.	Two or more cases in the media or credible online sources where open data is cited to have allowed marginalised groups to participate in either policy making, or accessing government services.	No evident impact at all

SCORING GUIDANCE

All societies have certain groups who are marginalised. This may be on grounds of age, gender, race, tribe, caste, class, disability, geographic location, and levels of poverty. Whilst these groups are not prohibited by licenses or technical mechanisms from accessing and using open data, they may not always be able to have effective access to open data. It has been argued however that open data can lead to more inclusive policy making and government services.

This may happen through the direct use of open data by marginalised groups, or through the work of intermediary organisations who support marginalised groups to access and use data, or who use data to campaign for the greater inclusion of marginalised groups in decision making or in receiving the benefits of public services.

SOURCE GUIDANCE

- Conversations with open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media;
- Blog posts and forum discussions;

I5: To what extent has open data had a noticeable positive impact on the economy?

SCORING CRITERIA

10	5	0
Clear rigorous evidence of a contribution of open data to a range of forms of new economic activity and/or economic growth; for example, a peer-reviewed study showing a positive impact on economic growth across a number of sectors.	Multiple cases in the media or credible online sources where open data is cited to have contributed to economic growth in at least two different sectors.	No evident impact at all

SCORING GUIDANCE

Open data may impact on the economy in a number of ways. For example:

- Through supporting the creation of new businesses based on open data;
- Through supporting existing businesses to lower their costs or become more efficient (for example, using weather or transport data to better plan their operations)
- Through supporting better economic planning;

You can read about a range of open data business models at:

<http://radar.oreilly.com/2013/01/open-data-business-models-deloitte-insight.html>

To score 10 the studies or credible sources cited should be about actual economic growth, rather than forecasts of potential economic impacts from open data.

SOURCE GUIDANCE

- Conversations with businesses, open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media;
- Blog posts and forum discussions;

I6: To what extent are entrepreneurs successfully using open data to build new businesses in the country?

SCORING CRITERIA

10	5	0
There are five or more examples of successful and thriving new businesses based around open data, employing staff and/or attracting significant external investment.	There are small scale examples of commercial apps, websites or other businesses built with open data. These remain niche or small scale businesses.	There is no evidence of entrepreneurial use of open data.

SCORING GUIDANCE

This question focuses specifically on entrepreneurial use of open data. See I5 for details of business models around open data.

For this question you may consider where national businesses are also drawing on open data from outside the country, from academia and from community sources, as well as open data published by the national government.

For example, a business created in the country based on the community-created [Open Street Map data](#) would count towards the examples for this question.

Be careful to distinguish business based on **open data** (explicitly citing certain open datasets as a key input into their work) from businesses based around open source software. Open source software is distinct from open data, and open source businesses do not count towards scores for this question.

SOURCE GUIDANCE

- Conversations with businesses, open data experts, government officials, NGOs and media;
- Reports and case studies from government departments, international organisations and NGOs;
- Media reports from domestic or international media;
- Blog posts and forum discussions;

Dataset questions

This section consists of a list of 15 datasets. For each dataset you are asked to complete a 10-point checklist to determine whether it is available as open data. The following pages detail how to answer each checklist question.

Is it open? Identifying whether a dataset is open

A dataset is considered fully open when it is freely accessible, machine-readable, and there are no legal barriers to re-use. However, there are a number of features of datasets that may make them more or less open, and that may be important to allow sustainable re-use. This 10-question framework steps through important features that factor in judging the openness of a dataset.

For the dataset you are assessing, answer the following questions Yes or No to the best of your knowledge and research. Research tips are provided for some questions.

The survey form includes details of specific evidence you should provide for each answer, and provides space for further justifications and source information against your assessments.

a) Does the data exist?

If the data is not collected in this country at all, then answer no to this question. However, if the data is collected at the level of government you are answering for in any form (e.g. paper or digital), then you should answer yes.

Examples:

- If you are answering for 'census data' and a census is carried out, with records kept on paper, then answer **yes**.
- If you are answering for 'census data' and no data is collected, or the data is only collected in one or two regions, answer **no**.

b) Is it available online from government in any form?

Can you find the data online in any digital format. This might include tables on web pages, PDF files, or scanned copies of paper documents, as well as other formats like Excel.

You should only answer yes if a substantial proportion of the data itself is available. Press releases or short summaries selecting just a few of the possible data points do not count. Under each particular dataset guidance is provided on what might constitute adequate online data.

Note: Ensure the source you find is the responsibility of government. A government domain name, or clear statement that the website is sponsored by government may provide evidence of

this. If you find a dataset available on a non-government website then further searching to find if a government source is also available may be required. If no other government source is available, answer **no** to this question, **but** continue to the following questions based on the non-government source you have located.

Examples:

- If looking for mapping data and the government has published image files containing maps with boundaries and locations of key infrastructure (e.g. roads, rail) then answer **yes**
- If looking at census data and the only data available is in PDF reports with tables and graphs, but without in-depth statistical tables covering the majority of main census indicators, then answer **no**

Take a note of:

Where the data can be found:

Any limitations on the dataset that is provided (e.g. if it is only a partial extract):

Search tips: locating datasets online

1. If the government has a central open data portal search here first

If you are not sure if the country has an open data portal search the web for relevant mentions. You might also find data portals from specific departments, particularly national statistics agencies.

You can also find a list of known Open Government Data portals at <http://census.okfn.org/catalogs/> and the Reneasler Polytechnic in the United States have brought together the contents of over 40 open data catalogues at http://logd.tw.rpi.edu/demo/international_dataset_catalog_search

Note: neither of these sources are authoritative, and may contain outdated data, or may miss out on data, particularly information not accessible in English. You should confirm for yourself any information on datasets gathered from these sites.

2. Check the websites of any government departments or agencies who may be responsible for this dataset - use their internal search features to look for the relevant data or to search for 'open data' to see if they contain a listing of available datasets.

Consider using the 'site:' and 'filetype:' operators in Google search to probe the websites if their in-built search is not very effective.

3. Search the web for mentions of this data or any other pointers to a possible dataset

In addition, to identify whether a dataset is available, and available as open data, you might:

- Enquire with the relevant government department directly;
- Ask on the mailing lists of the local open data community if one exists;
- Consult academics who work with this kind of data to identify whether they have found sources for it;

Using 'site:' and 'filetype:' operators.

Google search has two advanced search options that can help you locate data.

The 'site:' operator restricts search to files indexed by Google on the site(s) you specify.

For example, having located the website of the Geospatial agency in Indonesia I can run the search:

[site:bakosurtanal.go.id "open data"](https://www.google.com/search?q=site:bakosurtanal.go.id+open+data)

to search for mentions of the phrase "open data" anywhere on the website. I could also widen the search to check on any government website using 'go.id' as the site. For example:

[site:go.id "open data"](https://www.google.com/search?q=site:go.id+open+data)

Using the 'filetype:' operator I can look for files of a particular kind. For example:

[site:bps.go.id filetype:csv](https://www.google.com/search?q=site:bps.go.id+filetype:csv)

searches the Statistics Indonesia website for any CSV files that may contain relevant statistical data. To explore whether a particular dataset might be available I could guess at something I would expect to find within it and add this to my query.

c) Is the dataset provided in machine readable formats?

Can you find a copy of the data which can be opened in appropriate data manipulation and analysis software.

If there are multiple datasets that you could answer with respect to (for example, when a government publishes different data products based upon a census or land registry, or when

there are a number of different government service directories maintained) you should answer with respect to the best example you can locate in terms of availability of open data. For example, if the government holds directories on hospitals, community health centres and doctors, but only the list of hospitals is in a machine readable form, you should note this in the justification and continue to answer the questions with respect to this hospitals datasets.

You may need to search the site where you find the data for a machine readable form. For example, if on a data portal, go back to the root of the portal and search for all mentions of this dataset to find different versions. If on a government website, use the 'site:' operator to search the site for 'open data', 'datasets' or files of csv, xls or other types.

Examples

- If the data is only available as scanned copies of printed out documents in PDF or in image formats answer **no**. In these cases the only way to get data that can be sorted or manipulated would be to retype it into a spreadsheet. This shows the data was not originally machine readable.
- If looking at statistical data, and there is a version of the dataset you can load into spreadsheet software and sort and filter, answer **yes**
- If looking at statistical data and the only data available is in PDF files, or on web pages where copying and pasting the data into a spreadsheet does not let you easily sort and filter it, answer **no**
- If looking for mapping data and there are shapefiles (SHP), Points of Interest (KML) or other geographic data formats provided, answer **yes**
- If looking for mapping data and the only data available is in the form of images, even if in 'slippy maps' online, answer **no**
- If there is an API (Application Programming Interface) for the data, or bulk downloads provided in XML, CSV or JSON formats, answer **yes**

Take a note of:

Formats the data is provided in:

d) Is the machine readable data available in bulk?

Bulk access to data files allows developers and analysts to more flexibly build upon the data and to integrate it into other products, services and activities.

Sometimes it is only possible to access small extracts of a dataset using an online interface, as when, for example, you can only get a spreadsheet download after you build a query onto the dataset, or when data is scattered across tables or a very large number of download files on different web pages. Look for a link to bulk downloads. These might be organised by month for time-series data, or broken down into sub-files for very large datasets.

However, to qualify as bulk data it should be possible for a developer to easily collect a list of

download file URLs and then grab this data and feed it into an application.

Examples:

- If you are looking for crime data and the relevant agency publishes a list of monthly crime statistics files for download on one web page answer **yes**
- If looking for transport timetable data and the only way to get at a timetable is to fill in an online form for each transport route you are interested in, answer **no**

Take a note of:

Where bulk downloads can be found:

e) Is the dataset available free of charge?

If a bulk dataset is available, answer with respect to this. If the data is available online, but not in bulk, answer with respect to the online accessible version of the data.

Sometimes datasets are only available for a charge - either a one-off payment or a subscription fee. For example, when maps are licensed by the mapping agency and only provided for a fee, or when a country's legislation is compiled and provided for sale through a private company.

If a dataset is only available for a fee from a private company then please include notes on this in the comments.

Examples

- If looking for school performance data and the data is published for free in bulk on a government agency website, answer **yes**
- If looking for legislation and a company sells bulk access to a database of laws on CD for a price, and the data is not otherwise freely available, answer **no**

Take a note of:

Any charging regimes for the data:

f) Is the data openly licensed?

If a bulk dataset is available, answer with respect to this. If the data is available online, but not in bulk, answer with respect to the online accessible version of the data.

Unless the license that a dataset is provided under is explicitly stated, users of the data cannot be sure whether they have permission to re-use it, and under what conditions. For example, government datasets may be subject to copyright or intellectual property restrictions that make it illegal for a campaign group or company to build on top of the data or redistribute it.

Look for a clear licenses statement setting out the permission that a user has. This will only

count as an open license if:

- It clearly states that anyone has permission to re-use it;
- It does not restrict who can re-use the data (for example, through non-commercial restrictions);

Refer to the [Open Definition](#) and the list of [conformant licenses](#) for a detailed overview of what counts as an open license.

Examples:

- If there is no statement concerning how the data can be re-used, either next to the download, in the terms and conditions of the website, or on other relevant pages where the dataset can be found, answer **no**
- If there is a copyright or license statement that states the dataset can only be used for non-commercial use, or that it is only for 'personal' or 'research' use then answer **no**
- If there is a clear open data license statement associated with the dataset such as [Creative Commons Zero](#), [Open Database License](#), or [Open Government License](#) download answer **yes**
- If a statement of general permissions is associated with the dataset, and the intent of the statement is clearly and unambiguously intended to permit re-use of the data for all purposes, including commercial use, answer **yes**
- If there is statement of permissions, but it is ambiguous, and no other license is given, answer **no**

Take a note of:

The license the data is under:

g) Is the dataset up to date?

Many datasets are only useful if they are recent and kept updated.

Consider how often you would expect data of this form to be updated, and look for evidence that this is a recent version of the data.

Examples

- If looking at a list of public services, and there is a bulk data file for download, but it appears to have been generated once-off, and not updated since, answer **no**
- If looking at census data, and there are files from the last census, answer **yes**
- If looking at budget data, and there are data files from the most recent budget, answer **yes**

Take a note of:

The last update of the data was:

h) Is the publication of this dataset sustainable?

Sometimes open data appears to involve just one-off publication of datasets, without any process in place to keep the data updated.

Look for evidence that this was a once-off publication, or for evidence that the dataset is being kept regularly updated. Offer your best judgement as to whether it is likely that this open data will be kept updated for at least a further 12 months (or longer if the update cycle for this data is longer than 12-months as in the case, for example, of election data, where you would judge on the basis of whether it is likely that open data will be published on results at the next election).

In some cases the evidence available to make a judgement will be limited. In these cases, where you cannot locate any evidence that the data will be updated regularly, you can answer **no**.

Examples:

- If this is a type of data set that needs regular updates, but only one version has been produced, answer **no**
- If all the datasets on a data portal were added on one data over 6 months ago, and there has been no activity since, answer **no**
- If there appear to be regular updates to the dataset, for example, monthly versions posted on the web page, answer **yes**
- If there is a clear commitment from the responsible agency to keep their open data updated, answer **yes**

i) Was it easy to find information on this dataset?

Thinking particularly about questions 2a - 2f, was it easy to find information on this dataset.

Answer **yes** if you think a regular Internet user with a degree level education would be able to locate and find out about this dataset if they were looking for it. Answer **no** if you struggled to locate the dataset, or if you think a degree educated Internet user in your country would face considerable struggles in trying to find and find out about this dataset.

Examples:

- If once you located the dataset all the information you needed about it's status as open data was just a few clicks away, answer **yes**
- If you had to repeatedly go back to search engines to find out about dataset licenses etc, or if you had to draw upon privileged information or contacts to find out about the dataset, answer **no**

j) Are data URIs provided for key elements in the dataset?

URI stands for Uniform Resource Identifier. For the purpose of this question you can treat these the same as URLs (Uniform Resource *Locators*), or, as they are more commonly known, web

addresses.

Datasets contain many different entities. For example, a dataset on education might include record on about schools, regions and types of education programme. Advanced publication of these as open data, following [linked data principles](#) would involve creating a URI/web address for each of these entities, such that when a human or machine looked up that address they would be able to find data about that entity. The advantage of this is that when different datasets use the same URIs to identify a thing then connections can be more easily made between those datasets.

Checking if data URIs are provided may involve some detective work.

First, check to see if there is a 'Linked Data' version of the dataset is advertised. This may be provide in the 'RDF (Resource Descriptor Format)' file format, or as a file with the .ttl, .rdf, or .n3 file extension.

To answer **yes** to this question these formats are *not* required, but if you find these formats then it is likely that the answer is yes.

Another clue that linked data version of this dataset are available is when a 'SPARQL endpoint' is mentioned. SPARQL is a query language for linked data. Endpoints let people run their own queries against the data.

Secondly, if you don't find these formats apparent, look to see if there are unique web pages provided for all of the key entities in the dataset *provided by* the organisation or agency responsible for the dataset (Note: third party publication of URIs does not count, unless under agreement with the government organisation/agency).

For example, the UK Legislation website at <http://legislation.gov.uk> provides unique URIs for every bit of UK legislation, such as <http://www.legislation.gov.uk/ukpga/2012/9/contents/enacted>

If you find URIs about key entities from the dataset, then you can check to see if a machine-readable data representation of the entity is available. Look for links to RDF, XML, JSON or other data formats on the page.

You may also 'view source' of the page, and check for <link rel="alternate" type="X" ...> links in the header which have formats such as application/xml, application/json, text/csv or similar.

Examples:

- If queries for the dataset name and RDF, Linked Data and SPARQL provides no clear pointers to available data, answer **no**
- If you find URIs for key entities from the dataset, maintained by the agency responsible

for the dataset, and any machine-readable kind of data is available from these pages (CSV, JSON, XML etc.) answer **yes**

- If you find a series of web pages for each entry in the dataset (e.g. each school; each road) and this has a link to a .rdf version, which launches a file download, answer **yes**
- If you search and find a SPARQL endpoint which claims to contain the dataset, and this appears to be maintained by government, or under agreement with government, answer **yes**
- If you search and find an RDF version of the dataset, or a SPARQL endpoint which claims to contain the dataset, but this appears to have been built by a third party unrelated with government, answer **no** (this is because in these cases the URIs given to entities are not likely to be being maintained by government).

Extra resources on Linked Data:

- Wikipedia page on Linked Data: http://en.wikipedia.org/wiki/Linked_data
- Cool URIs: <http://www.digoiduna.eu/documentation/linkeddata-approach>

Take a note of:

Where the linked data is published:

Dataset details

This section details the 15 kinds of data you are asked to carry out research for. Given that countries vary in how they might collect or manage this data, rather than providing formal definitions of these datasets we give a range of examples of each. NOTE: ***The examples given would not necessarily score very highly on the 10-point scale, although most would score YES to at least questions (a) and (b).***

D1: Map data (full map coverage of the country)

A detailed digital map of the country provided by a national mapping agency and kept updated with key features such as official administrative borders, roads and other important infrastructure.

Examples:

The United Kingdom mapping agency, the Ordnance Survey, provides open data mapping products for download:

<http://www.ordnancesurvey.co.uk/oswebsite/opendata/discover.html>

The United States Census.gov site provides detailed mapping products for download:

<http://www.census.gov/geo/maps-data/>

The Mozambique mapping agency at

http://www.cenacarta.com/modules.php?name=Downloads&d_op=viewdownload&cid=5
provide shape files for download.

The existence of a country on **Google Maps** or other online mapping products does not count - as (a) this data is usually copyright; and (b) the mapping is often based upon data collected by government in the first place.

D2: Land ownership data

A dataset that provides national level information on land ownership. This will usually be held by a land registration agency, and usually relies on the existence of a national land registration database.

Open data published from this dataset might take the form of:

- **Full details of land ownership** - this is the ideal case, but often whilst governments may hold a full database of land ownership, they will only release as open data some subsets of this, such as those listed below:
- **Datasets on government owned land**
- **Data on land transactions (e.g. change of ownership, or prices paid for land)**

Note what kinds of data your assessment covers, if any.

Examples:

The New Zealand Land Registry makes data on 'land parcels' available as open data as listed at <http://data.govt.nz/dataset/show/1333>

The UK Land Registry publishes monthly residential sales 'Price Paid' data on land transactions at <http://www.landregistry.gov.uk/market-trend-data/public-data/price-paid-data>

This [news story](#) reveals that Delhi has implemented a GIS system for land ownership, but also implies that a national land ownership database does not currently exist.

D3: Government service directory

This is a dataset providing details of the key services provided by government and contact details that citizens can use to secure these services. Generally this will need to be at a granular level in order for citizens to access useful contact details, such as the number of a local police station, or the contact details of the local schools.

Sometimes the components of service directory will be scattered across different departments and agencies. If you can locate datasets that provide, at a minimum, a service name, address or location, and contact details, for three of the following services, then consider that this dataset exists: health services, schools, police services, government service centres, libraries and social services.

You may need to speak with officials or expert informants to identify whether government holds this information internally, or to identify when such data is not collected at all at a national level.

Examples:

The United Kingdom Local Direct Gov Service List provides links to local authority websites for a wide range of government services, [and is published as open data](#).

The [Kenya Data Portal](#) lists a number of datasets on government services, but there are only clear datasets on health services and school locations, and these do not contain contact details.

D4: Detailed census data

Countries may release detailed statistical data from a national census through their National Statistics office, or in collaboration with an agency who was supported their census work. A detailed census data should generally contain information such as age, gender and education levels broken down at least regionally. Usually the data you will be able to access will be extracts

from the census, rather than the ‘raw’ data, which includes personal information that cannot be released as open data.

The Wikipedia list of government’s conducting a census may be useful:

http://en.wikipedia.org/wiki/Population_and_housing_censuses_by_country

Examples

The [Moldovan Open Data Portal](#) contains datasets on population distribution drawn from the Census.

[Detailed bulk data tables](#) from the UKs 2011 Census are available online.

[Statistics Canada](#) provide in depth census tables for download.

D5: Detailed government budget

The budget sets out government spending plans.

Research tips:

The Open Budget Survey (not universal coverage) regularly assesses whether budget data is available from a government <http://survey.internationalbudget.org/>.

The OpenSpending.org platform have brought together many budget datasets, and may help you locate whether the government is publishing open budget data or not. In some cases data is being made available on open spending by third-parties converting from printed or PDF budget records. In these cases you can judge that data is available, but should decide whether or not this is sustainable, given the effort required on the part of civil society.

Examples:

South Africa publishes detailed information on the countries budget online at <http://www.treasury.gov.za/documents/national%20budget/default.aspx> - a search on the site for xls files (using google> ‘site:treasury.gov.za filetype:xls budget’) suggests that datasets for elements of the budget may also be available, although further investigation is needed to confirm.

The UK Data Portal lists the PESA budget dataset at <http://data.gov.uk/dataset/hmt-pesa09-chapter1>

D6: Detailed data on government spend

Government spending data details where money has actually been spent, and usually contains listings of transactions to various suppliers.

Examples

The Uruguay Office of Planning and Budget publish a visualisation of public spending, along with underlying datasets: <http://agev.opp.gub.uy/advni/>

The UK has mandated detailed publication of all government spend over £25,000 and over 500 spend datasets can be found at data.gov.uk:
<http://data.gov.uk/data/search?tags=spend-transactions>

D7: Company register

A company register should contain details of the registered firms within a country. Many registers provide online databases, but these are often only accessible for searching on a company-by-company basis, or behind a log-in. Some sell access to company data, but do not provided any open data.

In some countries company registration takes place at a state or regional level. In these cases, check for a national aggregation of company registers.

Research tips

The Wikipedia list of company registrars by country may prove useful in research:
http://en.wikipedia.org/wiki/List_of_company_registers

OpenCorporates.com have scraped data from many different corporate registries. Each record should link back to it's source, pointing you to details of the government held data. If the data only appears to be available in bulk form on Open Corporates, and is not available in bulk from the government, then do not award the dataset a mark for sustainability.

Examples:

[UK Companies House](http://UKCompaniesHouse) list a 'free public data product' on their company information pages, which consists of a download of basic company information.

The Slovenian government provide a list of taxpayers which provides registration numbers for legal persons (i.e. companies) as a data dump. The license of the data is unclear.

D8: Legislation

The laws of a country may be published online. In some countries legislation has been made available in structured formats as open data, whereas in others bulk access to legislation is only available for a fee.

Research tips:

Legislative open data will often be provided using eXtensible Markup Language (XML), so

searching for '[country] legislation XML' may help locate open legislation data.

Examples

The UK [Legislation.gov.uk](http://legislation.gov.uk) site provides laws as structured data as described in this [Legislation as open data](#) presentation. The site provides stable URLs for laws.

XML data files of recent US Laws are available through the Library of Congress Thomas site, as detailed at <http://xml.house.gov/>.

D9: Public transport timetables

A timetable provides details of when and where public transport services such as busses and rail services are expected to run.

A number of countries have national transport information systems that can provide detailed open data on public transportation. In other countries this information only exists at the city level, so for these countries you should carry out an assessment for the capital city, and note this in supporting information.

Research tips:

Sometimes transport information is provided through an 'API' (Application Programming Interface) rather than as bulk downloads.

The GTFS data exchange provides details of many local areas sharing public transport data <http://www.gtfs-data-exchange.com/agencies/bylocation>

Examples:

[The Traveline National Dataset of public transport timetables](#) provides UK public transport timetables.

Labs Ruter provide an [API onto real-time transport information for Norway](#), released under an open license.

D10: International trade data

International trade data records may include details of the import and export of specific commodities, or balance of trade data against other countries.

Examples:

Data.gov.uk lists [a range of trade statistics available](#) under the Open Government License.

The Central Statistics Organisation of Afghanistan publish [a list of exports by commodity](#)

[and country](#) in a Excel spreadsheet on their website, along with other trade statistics.

D11: Health sector performance

The performance of health services in a country has a significant impact on the welfare of citizens. Look for ongoing statistics generated from administrative data that could be used to indicate performance of specific services, or the healthcare system as a whole.

Health performance data might include:

- Levels of vaccination;
- Levels of access to health care;
- Health care outcomes for particular groups;
- Patient satisfaction with health services;

Where health sector performance data is split across many different datasets, base your evaluation on the best case you find amongst the first 5 datasets you find.

Research tips:

[The Global Health Data Exchange](#) maintain country profiles and a repository of datasets for countries across the world. The [advanced search](#) can be used to find 'Administrative Record' data types. This can help you locate national sources of health data, and explore whether these are open datasets. Not all the sources listed on the Exchange are from national governments, so check carefully when exploring available data.

National ministry of health websites are a useful starting point for a health data search in countries that do not have a national open data portal.

Examples:

The [UK Department of Health](#) publish extensive statistics on the performance of the health service.

The [Kenya Open Data Portal includes](#) statistics on levels of immunisation in the country, although the data is not updated.

D12: Primary or secondary education performance data

The performance of education services in a country has a significant impact on the welfare of citizens. Look for ongoing statistics generated from administrative data that could be used to indicate performance of specific services, or the education system as a whole.

Performance data might include:

- Test scores for pupils in national examinations;
- School attendance rates;

- Teacher attendance rates

Simple lists of schools do not qualify as education performance data.

Examples:

The [Kenya Open Data Portal](#) includes a range of education datasets, including attendance and test score statistics.

Brazil [publishes a microdata census of schools annually](#).

[The New Zealand Open Data Portal](#) lists a range of education and learning outcomes datasets, including primary and secondary school achievements.

D13: Crime statistics

Crime statistics can be provided at a variety of levels of granularity, from annual returns on levels of crime, to detailed real-time crime-by-crime reports published online and geolocated, allowing the creation of crime maps.

Research tips:

Look for data from a national criminal records agency, or policing department.

Examples:

The Indian [National Crime Records Bureau](#) publish annual crime and prison statistics.

The [UK Police Force publishes street level](#) crime data on a monthly basis, sharing the location of individual crime reports.

D14: National environment statistics

Countries may keep record of a number of different environmental statistics. Check for data on carbon emissions, levels of pollution of water or air, and deforestation. Base your assessment on the best available dataset from this set.

Environment statistics may be available from research institutions. In these cases only judge these as open government data when it is clear that there is national government support for their collection and ongoing management.

Examples:

[The UK Department for Food and Rural Affairs](#) maintain an archive of air quality information.

The [Mexican 'Your Government Maps' service](#) provides maps on a variety of environmental indicators.

D15: National election results

Election results report details of who received votes in elections to a national parliament or assembly.

Election results are often published online as PDF documents or on web pages, or may be made available in machine readable forms. Often this data will be published by an elections commission, and republished by the media. Look for the original source where possible.

Examples:

The [Swedish Election Authority](#) provide detailed downloads of election information, returns and results.

The [UK Electoral Commission](#) publish a spreadsheet of recent election results.