

Open Data Publication Plan



Except for images or where otherwise stated this document is © National Library of Scotland and is licensed under the Creative Commons Attribution 4.0 International Licence. To view a copy of this licence, visit: <http://creativecommons.org/licenses/by/4.0/>

Cover image: opensource.com, <https://flic.kr/p/aZhtF4>. Licensed under Creative Commons Attribution Share-Alike 2.0 Generic Licence. To view a copy of this licence, visit: <https://creativecommons.org/licenses/by-sa/2.0/>

1. Summary

- We will publish our non-personal and non-commercially sensitive data as 3-star¹ open data.
- We will use non-proprietary formats and licences that are conformant to the Open Definition.²
- We will continue to publish datasets and add these to our Open Data Register as they are identified.
- We will make our data available through our dedicated data publication website, the Data Foundry.

2. Open data vision

The National Library of Scotland recognises and welcomes the societal, economic, governance and institutional benefits of open data. The opening of public data aligns with our mission to ‘make a significant and lasting contribution to global knowledge and the memory of the world’. We believe that opening our public data will further our strategic vision to be one of the leading national libraries in Europe.³

3. Goals

The Scottish Government has asked public organisations in Scotland to prepare plans for publishing their open data.⁴ We have separately identified two long-term goals for our plan that are specific to the National Library.

By opening our data we aim to generate:

1. **Better data:** We are expert in what we do, but we recognise that there will always be ways to improve our data and in turn our business and our public service. We aim to find improvements in our own management, maintenance and use of our data by implementing this plan.
2. **Wider re-use of our data:** We create our data for specific purposes, to achieve business and regulatory requirements. Some of our data is already shared in pursuit of those ends, for example for the creation of union catalogues. However, we recognise that data can be applied to further ends, which are often not imaged by data creators, to wide benefit. We aim to encourage wider re-use of the data that we create through implementation of this plan.

¹ Using Sir Tim Berners-Lee’s five-star linked open data deployment model:

<http://www.w3.org/DesignIssues/LinkedData.html#fivestar>

² The Open Definition, <http://opendefinition.org/od/2.1/en/>

³ *The Way Forward*, National Library of Scotland strategy 2015-2020, <http://www.nls.uk/media/1194639/2015-strategy-2.pdf>

⁴ Open Data Strategy for Scotland, <http://www.gov.scot/Publications/2015/02/6614>

4. Plan

4.1. Data

We will publish our non-personal and non-commercially sensitive data as open data in accordance with the Open Definition. We already make data available, for example on our websites, in our catalogues, and through third parties such as The European Library.⁵ However, we have not published this data as open data.

We therefore have two priorities under this plan:

- 1. To publish as 3-star open data¹ the data that we already make available**
and
- 2. To identify our non-personal and non-sensitive data that we do not currently make available, and to publish this data as 3-star open data.**

We will identify datasets and data sources that can be published under this plan. We will add these datasets to the Open Data Register as they are published.

Certain datasets can be published more quickly than others. Some datasets only need to be licensed and encoded in an open format, while other datasets may need to be 'cleaned' to remove personal data or data contributed from other sources that we don't have permission to openly license.

In general, the data that we create and publish falls into two categories:

1. Collections as data
2. Data about our business

We will provide links to the data from relevant locations on our websites (e.g. we will link to open data related to our spending from our Freedom of Information (Scotland) Act publication scheme).

4.2. Licences and rights statements

We publish our data using rights statements or open licences that are open or conformant⁶ to the principles of the Open Definition. We will primarily publish our data using the following statements or licences:

⁵ The European Library, <http://www.theeuropeanlibrary.org/tel4/contributor/P02009>

⁶ Open Knowledge Foundation, <http://opendefinition.org/licenses/>

Licence	Description	Link
Public Domain Mark	This statement means that the associated data is free of known copyright restrictions and is in the public domain.	https://creativecommons.org/publicdomain/mark/1.0/
No Known Copyright	This statement means that the associated data is likely to be free of known copyright restrictions, but that we have insufficient information to confirm that the material is in the public domain.	http://rightsstatements.org/vocab/NKC/1.0/
Copyright Undetermined	This statement means that the copyright status of the associated data has been evaluated, but this process was inconclusive. The associated data may be protected by copyright.	http://rightsstatements.org/vocab/UND/1.0/
CC0	The Creative Commons CC0 'No Rights Reserved' statement waives all rights in content, meaning that no legal restrictions are placed on the use and re-use of the content.	https://creativecommons.org/publicdomain/zero/1.0/
CC-BY	The Creative Commons Attribution (CC-BY) licence allows the free use and re-use of content for commercial or non-commercial purposes, subject only to the requirement to credit the source of the content.	https://creativecommons.org/licenses/by/4.0/

The Scottish Government's Open Data Strategy calls on organisations to publish data under the Open Government Licence (OGL), which is maintained by The National Archives.⁷ We have decided to license our data with Creative Commons licences instead of the OGL for two reasons:

- The OGL was designed for use with Crown copyright material, and the Library is not a Crown copyright body
- The OGL is specific to the United Kingdom, whereas Creative Commons licences are used globally

⁷ Open Government Licence, <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>

However, the terms of the OGL and the CC-BY licence are compatible. Because we will primarily use the CC-BY licence or the more open CC0 mark, we are confident that our open data licences conform to the Scottish Government's strategic objectives for open data.

4.3.Formats

We will publish our data as machine-readable structured data using non-proprietary formats. Although we may use other open formats as required in respect of any particular set of data, we will primarily publish our data using the following formats:

Format	Description	Link
Plain text	Plain text files (.txt) contain only text, without formatting.	
CSV	Comma Separated Values (CSV) provides a file format for storing tabular data as plain text, using commas to separate fields. CSV is an open alternative to proprietary file formats (e.g. Excel).	https://tools.ietf.org/html/rfc4180
MARCXML	<p>Machine-Readable Categories (MARC) is a set of open standards for describing library items. The MARC standards are maintained by the Library of Congress.</p> <p>Extensible Markup Language (XML) is a markup language that allows data to be encoded in a format that is both human-readable and machine-readable. XML is defined by the W3C.</p> <p>MARCXML is a framework developed by the Library of Congress for using MARC data in the XML language.</p>	http://www.loc.gov/marc/ http://www.w3.org/TR/2008/REC-xml-20081126/ https://www.loc.gov/standards/marcxml/
ALTO XML	Analysed Layout and text Object (ALTO) is an open XML schema, maintained by Library of Congress.	https://www.loc.gov/standards/alto/
Dublin Core	Dublin Core is a metadata schema used to describe library collections, as well as other forms of digital and physical resources.	https://dublincore.org/

4.4. Level of openness

We will adopt Sir Tim Berners-Lee's 5-star deployment scheme for linked open data in assessing our openness. We will publish our data as 3-star open data or above, in accordance with the Scottish Government's Open Data Strategy. We will publish our data as 3-star open data by:

1. Making data available on the web with open licences (see 3.1 and 3.2)
2. Publishing machine-readable structured data (see 3.3)
3. Using non-proprietary formats (see 3.3)

4.5. Publication

We publish our open data for download on our Data Foundry website.

We will assess the necessity and benefit of making our data available on discovery sites or portals, for example at the UK or EU levels.

4.6. Timeline

We publish and update our data on a regular basis. On our Data Foundry website we normally indicate planned future releases of data.

5. Regulatory environment

The Scottish Government's Open Data Strategy (February 2015) calls on public organisations in Scotland to have Open Data publication plans in place and published online by December 2015.

The National Library of Scotland is subject to the Freedom of Information (Scotland) Act 2002 and the Environmental Information (Scotland) Regulations 2004, which require the making available of certain information. The Library is also subject to the Re-use of Public Sector Information Regulations 2015, which requires the making of public information useable and re-useable.

The European Commission's 2011 report on open data identified the opening of public information as an 'engine for innovation, growth and transparent governance'.⁸ In 2013, the members of the G8, including the United Kingdom, signed an Open Data Charter, which calls for 'Open Data by Default'.⁹

6. Roles and responsibilities

6.1. Overall responsibility

The Associate Director of Collections and Research and the Associate Director of Digital have overall responsibility for the Open Data Publication Plan.

⁸ European Commission, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>

⁹ Open Data Charter, <https://www.gov.uk/government/publications/open-data-charter>

6.2.Operational responsibility

The Rights and Information Manager and the Digital Scholarship Librarian have operational responsibility for the Open Data Publication Plan.

7. Related documents

- **Open Data Register:** The Open Data Register is a tabular list of the datasets that we have made open or plan to make open. The Register is a live document, and datasets will be added to it as they are published.

8. Review

This plan will be reviewed every two years.