



OPEN DATA POLICY

Joint Nature Conservation Committee (JNCC)

Open Data Policy v1

May 2018

About JNCC

JNCC is the public body that advises the UK Government and devolved administrations on UK-wide and international nature conservation.

Our work helps maintain and enrich biological diversity and conserve geological features. It also helps sustain natural systems, which provide the core “services” we all depend on like food, fresh water and clean air. In this way they contribute to economic growth and social well-being and are integral to sustainable development

Good policy-making, planning, development and risk management all depend on reliable, up-to-date information about biodiversity status and trends. Our role is to provide evidence, information and advice so decisions can be made that protect our natural resources and systems.

JNCC itself is a forum that brings together the UK’s four country conservation bodies. We advise Government and a wide range of bodies to help join up policy and to deliver a strong and cost-effective evidence base by helping to see that the best possible return is achieved from investment in research and surveillance in the UK and internationally.

Joint Nature Conservation Committee (JNCC)
Monkstone House, City Road, Peterborough, PE1 1JY
Telephone 01733 562626
Fax 01733 555948
Email comment@jncc.gov.uk
Website jncc.defra.gov.uk

JNCC SUPPORT CO. Registered in England and Wales. Company no. 05380206. Registered office as above.



JNCC © copyright 2018

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

Foreword

Open data is “data that anyone is free to access, use, modify, and share for any purpose – subject, at most, to measures that preserve provenance and openness” (The Open Definition).

Open data is recognised in government policy and by civil society as a system of principles and best practice designed to increase the accessibility and maximise re-use of useful data and drive benefits for society and the economy. Within biodiversity and conservation, open data complements other progressive movements such as open science and open access.

Last year JNCC and the UK’s statutory nature conservation bodies agreed a future ambition to improve the open availability of biodiversity data by 2020. In support of this ambition JNCC has developed a plan to release all its data at the level of detail originally captured, under the terms of an open licence, except where there are legitimate reasons not to publish.

All data collected under partnership will be released (possibly at a reduced level of detail) within two years, and made fully open within five years. Environmentally sensitive data will be made openly available to the same timescales at the highest level of detail consistent with the avoidance of harm.

Public authorities in the UK have a general obligation to proactively disseminate environmental information that they hold to the public by electronic means. This includes data from monitoring of activities that could affect the environment. As an advisory body to government on biodiversity and nature conservation, JNCC produces and maintains a large number of potentially useful datasets. JNCC's open data programme will increase public availability of environmental information and promote transparency in decision-making about the environment.

JNCC also aspires to be a pioneer, innovator and inspiration to others within the open data community by both internal action and, in cooperation with partner organisations, actively engaging with users to promote re-use of open data. JNCC will also work to significantly reduce its own reliance on non-open data.

JNCC’s Open Data Policy is central to this approach. Publication of open data is bounded by standards and principles that cover a range of subjects such as licensing, accessible formats, communications, project planning, and risk assessment. The JNCC policy draws together key material from these areas and presents it in a cohesive manner as rules and advice for implementation within an organisation. The policy includes a framework for risk management of open data, with a particular focus on protection of personal data and avoidance of ecological harm.

This policy has been written with JNCC staff as the primary audience but has also been designed for open publication so that it may be adapted for use by other organisations, either as policy or as a basis for guidance. The full text of JNCC’s Open Data Policy is re-usable under the Open Government Licence.

For more information about JNCC open data please visit our site at jncc.defra.gov.uk/opendata. If you have any questions or feedback about this policy or about JNCC’s data please contact data@jncc.gov.uk.

Contents

Open Data Policy	5
1 Purpose and coverage.....	5
2 Definitions and usage	5
3 Introduction.....	6
4 Legal context	7
5 Monitoring implementation.....	9
6 Licences	9
7 Formats and accessibility.....	10
8 Documentation	11
9 Publication platforms and channels	12
10 Access controls and charging	13
11 Procurement and project planning	14
12 Communications.....	15
13 Assessment and risk management.....	16
13.8 Personal data	16
13.9 Ecological harm and protection of the environment.....	17
13.10 Duty of confidence	17
13.11 Volunteered information (non-personal).....	17
13.12 Intellectual property	18
13.13 Risks that do not prevent publication as open data	18
13.14 Mitigation of risks	18
Annex A – Glossary	19
Annex B – Primer on intellectual property rights related to datasets.....	25
Annex C – Supporting information on the legal context to open data	26
Annex D – Assessment tools and marks of quality	29
Annex F – Supporting notes on attribution statements	32
Acknowledgements.....	33

Open Data Policy

1 Purpose and coverage

- 1.1 The aim of this policy is to set out JNCC's approach to publication of open data, including the adoption of technical standards, support for user needs, and engagement with the open data community.
- 1.2 Development of this policy supports JNCC strategic objectives such as embracing innovative approaches to data management and providing mechanisms for open and efficient sharing of environmental data.
- 1.3 This policy has been written foremost for the use of JNCC staff, but is designed to be published openly and to be suitable for adaptation by partner organisations or by any other organisation that wants to pursue a similar approach to open data.
- 1.4 Nothing in this policy constitutes terms for use or re-use of any specific data.
- 1.5 Compliance with this policy is compulsory for all JNCC staff. Staff involved in the production or management of JNCC data assets in particular should make reference to this policy.
- 1.6 Parts of this policy address potential risks for JNCC associated with the publication of open data.
- 1.7 If JNCC staff consider that anything in this policy is insufficiently clear or is not workable, staff should contact the Head of Digital and Data Solutions who will be responsible for providing clarification or making amendments.

2 Definitions and usage

- 2.1 For the purposes of this policy:

'open data' is defined as "data that anyone is free to access, use, modify, and share for any purpose – subject, at most, to measures that preserve provenance and openness",

'shared data' refers to data that is shared, either with particular individuals or organisations or for specific purposes or for public access, under terms and conditions that are not 'open', and

'closed data' refers to data that should not be shared and can only be accessed by its subject, owner or holder.

- 2.2 Other specific terms used in this policy are defined in a glossary in Annex A.

- 2.3 In this policy:

must and **must not** (in bold type) are used to indicate points that are essential requirements. Failure to implement one of these points will be contrary to this policy. Data published contrary to any of these points may not be compliant with the definition of open data.

should and **should not** (in bold type) are used to indicate points that are best practice, good practice, or advisory. Implementation of these points is desirable but failure to implement them will not be contrary to this policy. Data published contrary to any of these points may still be compliant with the policy's definition of open data.

3 Introduction

3.1 In October 2017 JNCC published the following Statement on Open Data:

“JNCC aspires to be a pioneer, innovator and inspiration to others within the open data community by both internal action and, on behalf of the statutory nature conservation bodies, by actively engaging with existing and potential users to encourage, support and maximise re-use of its data. JNCC will publish an open data policy in 2018 and adopt recognised tools such as the Open Data Maturity Model, ODI Certificates, and the 5 Stars of Openness to improve the quality and accessibility of its data.

“By 2020 JNCC will release all its data at the level of detail originally captured, under the terms of an open licence, except where there are legitimate reasons not to publish. All data collected under partnership will be released (possibly at a reduced level of detail) within two years, and fully open within five years. Environmentally sensitive data will be made openly available to the same timescales at the highest level of detail consistent with the avoidance of harm. In parallel, JNCC will work to significantly reduce its reliance on non-open data.”

3.2 Public authorities in the UK have a statutory obligation to proactively disseminate environmental information that they hold, including data from monitoring of activities that could affect the environment, to the public by electronic means. As an advisory body to government on biodiversity and nature conservation, JNCC produces and maintains a large number of useful datasets. JNCC's open data programme is intended to increase public availability of environmental information and promote transparency in decision-making about the environment.

3.3 In particular JNCC intends to publish as open data:

all biodiversity data wholly collected or funded or commissioned by JNCC, by 2020 or within no more than two years from the date of collection, at the level of detail originally captured, and

all biodiversity data collected or mobilised with funding from JNCC or with public funding through JNCC, by 2020 or within no more than two years from the date of collection, at a reduced level of detail and, within no more than five years from the date of collection, at the level of detail originally captured.

3.4 Open publication of biodiversity data as set out in 3.3 may be subject to legal constraints, and to redactions or exemptions required by JNCC's risk assessment process for open data.

3.5 The timescales set out in 3.3 may be extended, by exception, to accommodate embargos on publication of data underlying or in connection with release of reports, papers in journals, official statistics, submissions to international bodies, and similar arrangements.

3.6 JNCC recognises that open data exists as both a concept in government policy and as an agenda in civil society. JNCC's approach to open data will comply with all relevant legislative requirements and legal restrictions. JNCC will also endeavour, as far as practical, to align its approach to open data with external policy guidance on open data and with norms of good practice established within the wider open data community.

3.7 JNCC also recognises the principles and best practices for the release of governmental open data set out in the

international Open Data Charter,¹ to which the UK is a signatory.

3.8 In particular JNCC acknowledges the following as key features of open data. Open data

- **must** be re-usable under an open licence (or similar status) that allows use, redistribution, modification, and compilation for any purpose;
- **must** be provided as a whole and at no more than the marginal cost of reproduction, and **should** be downloadable via the Internet without charge;
- **must** be available in an open, accessible, and machine-readable format; and
- **must** be accompanied by any additional information necessary for licence compliance (such as attribution requirements).

3.9 JNCC also acknowledges the following as key principles of open data.

Government data **should** be open by default. There will be legitimate reasons why some data cannot be released as open data, but any such decision **should** be supported by a clear justification.

Open data **should** be supported by clear documentation that provides data users with sufficient information to understand the source/s, context, nature, and analytical limitations of the data.

Public bodies **should** maintain and publish accurate and up-to-date inventories of their data holdings.

Public bodies **should** actively facilitate and encourage re-use of the open data they publish.

4 Legal context

4.1 Licensing of open data depends on permissions granted by the organisation/s or individuals that hold intellectual property rights (IPR) in the data. Implementation of this policy therefore requires a basic understanding of copyright and related rights that apply to datasets. Please see Annex B for a primer on this subject.

4.2 In the UK most publication of open data by government is the product of policy initiatives rather than of any statutory requirement. However some specific releases of open data are either required by law or involve the application of open licensing and standards to data that is required by law to be available to the public.

4.3 JNCC staff **should** be aware in particular of the following legislation related to re-use of public sector information, access to information, data protection, and the creation of spatial data infrastructure, as these legal requirements are relevant to the publication of open data by the public sector. Please see Annex C for additional notes and links to more information about this legislation.

4.4 The Re-use of Public Sector Information Regulations 2015 (RPSI or RoPSI)² set out the legal framework under which UK public sector bodies are required to allow re-use of their information. These regulations implement European Union directive 2013/37/EU,³ known as the revised 'PSI Directive'.

¹ <https://opendatacharter.net/principles/>

² <http://www.legislation.gov.uk/uksi/2015/1415/contents/made>

³ <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013L0037&from=FR>

- 4.5 The Freedom of Information Act 2000⁴ and the Freedom of Information (Scotland) Act 2002⁵ (FOI) create a public right of access to information held by UK public authorities. This right is subject to a number of exemptions. Environmental information is exempt under FOI if access to that information is covered by separate environmental information regulations.
- 4.6 The Environmental Information Regulations 2004⁶ and the Environmental Information (Scotland) Regulations 2004⁷ (EIR) create a public right of access to environmental information held by UK public authorities.
- These regulations implement provisions of the Aarhus Convention and EU Directive 2003/4/EC⁸ on public access to environmental information, and require public authorities to proactively publish some environmental information including data.
- 4.7 JNCC staff **should** be aware of JNCC's Guidance: Access to Information. Due to the nature of JNCC's work, access requests for data held by JNCC are more likely to be handled under EIR than under FOI.
- 4.8 Any dataset that a public authority would be legally unable to release to the public in response to an access to information request will not be eligible for publication as open data.
- 4.9 The Statistics and Registration Service Act 2007⁹ sets out arrangements for the governance and management of official statistics across the UK. Guidance produced under these arrangements will influence standards for publication of some open data.
- 4.10 JNCC staff **should** be aware of the UK Government Security Classification Framework¹⁰ which classifies data according to its content and suggests appropriate security procedures. Data in some GSC categories will not be eligible for open publication.
- 4.11 A new legal framework for data protection, the General Data Protection Regulation (GDPR),¹¹ applies in the European Union including the UK from May 2018. GDPR is implemented in the UK by a Data Protection Act¹² that also extends to areas of data protection beyond the scope of GDPR.
- 4.12 Open datasets do not normally contain personal data. However recognition of personal data is essential to proper risk assessment of open data publication. JNCC's processes for publication of open data **must** comply with JNCC's Data Protection Policy.
- 4.13 INSPIRE¹³ is an EU directive that establishes a common framework for spatial data infrastructure for the purposes of EU environmental

⁴ <https://www.legislation.gov.uk/ukpga/2000/36/contents>

⁵ <http://www.legislation.gov.uk/asp/2002/13/contents>

⁶ <http://www.legislation.gov.uk/uksi/2004/3391/contents/made>

⁷ <https://www.legislation.gov.uk/ssi/2004/520/contents/made>

⁸ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:041:0026:0032:EN:PDF>

⁹ <https://bit.ly/2lohwc9>

¹⁰ <https://www.gov.uk/government/publications/government-security-classifications>

¹¹ <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr>

¹² <http://www.legislation.gov.uk/ukpga/2018/12/enacted>

¹³ <https://inspire.ec.europa.eu/>

policies. INSPIRE does not mandate publication of open data but public authorities that publish data to comply with INSPIRE may also publish that data under an open licence.

- 4.14 The Prime Minister's letter of December 2017¹⁴ provides the policy basis for departmental requirements on open publication of "transparency" datasets such as monthly transactions, contracts, and organograms.

5 Monitoring implementation

- 5.1 The Open Definition¹⁵ **should** be used as a reference for whether an individual published dataset meets the minimum criteria for compliant open data. A published dataset that is not compliant open data **should not** be referred to as open data.

- 5.2 The following assessment tools and marks of quality, described in more detail in Annex D, will be used to assess implementation of this policy as it applies to key datasets and progress at the organisational level.

The 5-star deployment scheme for Open Data,¹⁶ also known as the "five stars of openness", is a scale for describing presentation of open data and linked open data.

Open Data Certificates¹⁷ are a tool devised by the Open Data Institute for self-certification of individual open datasets based on different levels of good practice.

The Open Data Pathway¹⁸ is a tool developed by the Open Data Institute with Defra to enable organisations to assess their open data maturity.

- 5.3 Beginning at the end of the 2018/19 financial year JNCC will produce and make public an annual report on open data achievements and progress in implementing this policy.

6 Licences

- 6.1 Open data **must** be provided under an open licence, unless the data is free of copyright and other intellectual property rights. (Data that is free of intellectual property rights may be provided as open data without a licence.)

- 6.2 Any dataset published as open data **must not** be accompanied by any additional terms that contradict the terms of the open licence or conflict with the Open Definition.

- 6.3 Open data published by JNCC **must** be provided under a licence that is conformant with the Open Definition and suitable for application to data. A list of standard conformant licences is available on the Open Definition website:

<http://opendefinition.org/licenses/>

- 6.4 The licence under which specific JNCC open data is available at any given point in time **must not** vary based on the user or the purpose for which the data will be re-used.

- 6.5 By default JNCC open data **should** be provided under the latest version of

¹⁴ <https://bit.ly/2IHmfcn>

¹⁵ <http://opendefinition.org/>

¹⁶ <http://5stardata.info/en/>

¹⁷ <https://certificates.theodi.org/en/>

¹⁸ <http://pathway.theodi.org/>

the Open Government Licence¹⁹ (OGL). The current latest version of the OGL is Version 3.²⁰

- 6.6 One of the following persistent links **should** be used to link to the latest version of the OGL in documentation accompanying any dataset released under the OGL:
- <http://www.nationalarchives.gov.uk/doc/open-government-licence/>
- <http://reference.data.gov.uk/id/open-government-licence>
- 6.7 The OGL is intended for licensing of UK public sector information, including but not limited to data. The OGL does not cover any personal data in the information.
- 6.8 Under some circumstances, for example if a dataset is published on behalf of a non-governmental organisation or contains personal data, a different open licence may be more suitable than the OGL.
- 6.9 Open data is more useful when it can be easily combined with other open datasets. Proliferation of open licences is undesirable because it reduces the interoperability of open datasets. Open data **should** be provided under a standard well-known open licence, and **should not** be provided under an obscure or bespoke open licence unless that is the only available means of open licensing.
- 6.10 The Creative Commons Attribution 4.0 International License²¹ (CC BY 4.0) **should** be the first alternative choice of licence for open data published or distributed by JNCC if the OGL is not suitable for any reason.

6.11 An open dataset may include data that is subject to third party intellectual property rights, provided the open dataset is provided under a licence that is compatible with all terms required by the third party for re-use of their data. The open dataset **must not** be provided under a licence that is more permissive than the terms granted by the third party.

6.12 Published datasets **should not** contain subsets of data that are subject to different licences unless there is a good reason for presenting the data in this form. Any such presentation of data **must** clearly distinguish those licensing differences either in the dataset itself or in documentation published with the dataset.

7 Formats and accessibility

7.1 Open data **must** be provided in a form that is machine workable, and **must** be provided in an open format. An open dataset **must** be provided as a whole.

7.2 Publication of open data **should** conform to recognised open standards and **should** in particular conform to open standards adopted for use in government technology.²²

7.3 Open data **should** be provided in a form that is human readable.

7.4 Open data **must** be provided in at least one open format. A format is open if there are no monetary or other restrictions on its use and it can be fully processed with at least one software tool that is free of charge and generally available.

¹⁹ <http://www.nationalarchives.gov.uk/doc/open-government-licence/>

²⁰ <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>

²¹ <https://creativecommons.org/licenses/by/4.0/>

²² <https://www.gov.uk/government/collections/open-standards-for-government-data-and-technology>

- 7.5 Open data **should not** be released in formats that are obscure or that are supported only by software that is not readily discoverable.
- 7.6 Publication of open data in an open format may involve changes to the presentation of the dataset. However open data **should not** be provided in a format that requires a significant reduction in the content, granularity, or quality of the dataset.
- 7.7 Open data **must** be provided as a whole and **should** be provided as a single bulk download. For some large or complex datasets a small collection of downloads available from a single location will be an acceptable alternative to a single bulk download if this approach makes the dataset more accessible.
- 7.8 If the usability or accessibility of an open dataset is dependent on code or software that is not already freely available, the code or software **must** be provided with the open data under a suitable open licence. Code or software **should** be provided under an Open Source Initiative compatible licence ²³ such as the MIT licence. ²⁴
- 7.9 Release of open-licensed data through an open API is good practice for datasets where there is an estimation of sufficient demand or where the data is frequently updated. Open data **should** be supported by an API if there is a benefits case that justifies the necessary investment of resources.
- 7.10 A dataset provided only through an open API will not be open data if the user of the API cannot extract or download the whole dataset. A dataset provided only through an open API is also unlikely to be recognised as open data if the user

has to write code or learn a syntax for querying the API in order to obtain a copy of the whole dataset.

- 7.11 Please refer to Annex E for supporting notes on formats and accessibility.

8 Documentation

- 8.1 Open data **must** be accompanied by documentation that clearly sets out terms for re-use of the data, including any information necessary for a user to comply with those terms.
- 8.2 Open data that is provided under the OGL, or under any other licence that requires the user to acknowledge the source of the data with an attribution statement, **must** be accompanied by specification of an attribution statement.
- 8.3 Any specification of an attribution statement **must** identify the source of the dataset, and **must** incorporate any attribution statements specified by the terms of re-use of any third-party data included in the dataset. Please refer to Annex F for notes on writing attribution statements.
- 8.4 Open data **should** be accompanied by (or provided with a link to) metadata that conforms to a recognised schema, such as UK GEMINI ²⁵ for spatial data.
- 8.5 If an open dataset is part of a series or a version of a dataset that has been published previously, documentation provided with the dataset **must** be sufficient to distinguish the dataset from other datasets in the same series or other versions of the dataset.

²³ <https://opensource.org/licenses>

²⁴ <https://opensource.org/licenses/MIT>

²⁵ <https://www.agi.org.uk/agi-groups/standards-committee/uk-gemini>

- 8.6 Good design of a dataset can support understanding but datasets are rarely self-explanatory. Open data **should** be accompanied by documentation that is sufficient to enable a potential user to understand the dataset. Standard metadata with an abstract will fulfil this purpose for some but not all datasets.
- 8.7 The content and extent of documentation necessary to enable a user to understand a dataset will depend on the particulars of the dataset. Documentation **should not** make assumptions about the level of knowledge of the user or the purpose for which the open data will be re-used.
- 8.8 Documentation provided with open data **should** make use of existing material where available, including links to any related online resources. Documentation **should** be understandable by a general audience. Where the dataset is of a scientific or technical nature the documentation **should** focus on the dataset and **should** signpost rather than explain any domain knowledge necessary to understand the data or documentation in its full context.
- 8.9 If open data is published with any personal data or personally identifiable information (PII), or data that the publisher considers may be personal data or PII, whether or not that data is covered by an open licence, the data **must** be accompanied by an information warning or other documentation that draws the attention of the recipient to the personal nature of the data.
- 8.10 If open data is published with any third-party data that the publisher is not authorised to license, or is not authorised to license on the same terms, the data **must** be accompanied by an information warning or other documentation that clearly distinguishes those categories of data.
- 8.11 Open data **should** be accompanied by an information warning that highlights any significant concerns about misinterpretation or misuse of the data. An information warning may for example include caveats about the quality of the data or identify purposes for which the data is unsuitable.
- 8.12 Information warnings that accompany open data **must** be advisory. Information warnings and other documentation provided with open data **must not** place any mandatory constraints on use of the data that contradict the terms of the open licence or conflict with the Open Definition. However an information warning may draw the attention of users to existing legal restrictions on use of the data.
- 8.13 Open data that is published as a static dataset and has potential for citation in scientific or scholarly works **should** be allocated a digital object identifier (DOI).

9 Publication platforms and channels

- 9.1 Datasets created by JNCC that are eligible for publication as open data **must** be published on a JNCC domain (jncc.defra.gov.uk or jncc.gov.uk) or at an alternative online location under JNCC's administrative control, in addition to any publication elsewhere.
- 9.2 Specific datasets created by JNCC that are eligible for publication as open data **should** also be submitted for publication or archiving by any reputable organisation that maintains a suitable platform for that purpose.
- 9.3 If JNCC open data is submitted to a platform maintained by an external organisation the data **must not** be submitted on a basis that permits re-use under any licence that is more permissive than the licence already

applied to the dataset, unless this variation in terms has been approved through the risk assessment process for JNCC open data.

- 9.4 Preparation of JNCC open data for submission to a platform maintained by an external organisation **should** comply with that platform's requirements for format and documentation, but **must** otherwise follow this policy.
- 9.5 The OGL and other open licences permit redistribution and compilation of open data. Any licensee may republish JNCC's open data via other platforms and channels that are not controlled by JNCC provided the licensee complies with the open licence.
- 9.6 Open data published by JNCC **should** be recorded on Data.gov.uk (DGU), the UK's central online catalogue for government datasets. Any DGU record maintained by JNCC about a publicly accessible dataset **must** include either a direct link to the dataset or a link to a location where the dataset is readily discoverable.
- 9.7 Any DGU record maintained by JNCC **should** be updated promptly to reflect any material change to the location, availability, terms of use, or documentation of the dataset.
- 9.8 The availability and location of JNCC open data and other authoritative open data resources **should** be signposted in reports and other material published online, where they are relevant or related to the material.
- 9.9 JNCC open data and other authoritative open data resources **should** be accessible or signposted from any online application, search interface, mapping interface, or interactive visualisation that makes use of that data. Design of any such application, interface, or visualisation **should** facilitate access to and

discoverability of whole open datasets.

- 9.10 The terms of use for any platform through which JNCC open data is provided **should not** contradict the terms of the open licence or conflict with the Open Definition. If the terms of use for such a platform do conflict with the Open Definition, the dataset **must** be available from an alternative location on a basis that does conform to the Open Definition.

10 Access controls and charging

- 10.1 Open data **should** be downloadable via the Internet without any form of access control unless there is a compelling reason for that control. All forms of access control are potential barriers to the accessibility and re-use of open data.
- 10.2 Open data **must not** be subject to any form of access control that discriminates against any person or group, or that restricts anyone from making use of the data in a specific field of endeavour. Examples include platforms that limit access to subscribers and processes that only provide access to data on request for a limited range of purposes.
- 10.3 Any data collected via any form of access control for open data (such as registration, identification, or authentication of requestors, recipients, or users) **must** be managed in compliance with JNCC's Data Protection Policy.
- 10.4 Open data **should** be accessible without charge and **must not** be subject to any charge for access that is more than the marginal cost of reproduction of the data provided.
- 10.5 The marginal cost of reproduction is the cost of producing one more unit of a good. For open data and other

digital goods this will normally be very low and treated as nil for practical purposes. However there will be limited circumstances under which this cost **should** be charged for access to open data, as for example where the egress cost of supplying a copy of a very large open dataset cannot reasonably be funded by the publisher.

- 10.6 Charges **must not** be made for re-use of open data (as distinct from access to the data). Data that is re-usable under an open licence may be redistributed for a charge or included in commercial data products, but is not open data under those conditions.
- 10.7 Based on the principle of open by default JNCC **must not** make any charge for re-use of public task information that is eligible for publication as open data.
- 10.8 This policy does not place any restrictions on charging for access or re-use of shared data that has been collected or produced for a purpose other than delivery of a public task.

11 Procurement and project planning

- 11.1 Any new contract or agreement entered into by JNCC for procurement of the collection or creation of data **must** be compatible with the intentions on publication of biodiversity data in section 3.3 of this policy.
- 11.2 Any contract or agreement for procurement of the collection or creation of data entered into by JNCC prior to the implementation of this policy **should** be reviewed to assess its compatibility with the intentions in 3.3. Any such contract or agreement that conflicts with the intentions in 3.3 **should** be renegotiated to align with those intentions, provided renegotiation is feasible and

compatible with the aims of the contract or agreement.

- 11.3 Any collection of new data funded entirely by JNCC **must** be procured on a basis that either provides JNCC with rights to re-use the data under an open licence or obliges another party to publish the data as open data. This requirement applies only to new data where collection is funded entirely by JNCC and does not place any restriction on JNCC use of third-party data licensed on other terms.
- 11.4 Any new contract or agreement entered into by JNCC for procurement of the collection or creation of data **must** include a clear statement on ownership of any intellectual property created under that contract or agreement.
- 11.5 Some JNCC supply chains for data are complex and it is important that partners keep good records of their data sources. Any new contract or agreement entered into by JNCC for procurement of the collection or creation of data **must** include a requirement that any party procuring data from third parties will record evidence of the provenance and terms of re-use of that data and share those records with JNCC on request. Those records **should** include copies of any relevant licences and correspondence.
- 11.6 Any use of third-party data by JNCC to create either a dataset in which JNCC has intellectual property rights or a dataset that will be provided to or shared with any other party **must** be supported by records of the provenance and terms of re-use of the third-party data. This includes third-party data that is re-used under an open licence.
- 11.7 Data outputs, including the potential to release outputs as open data, **must** be considered at the planning and initiation stages of any new project undertaken or wholly funded by

JNCC. Project planning **should** follow approaches to procurement, and make use of software and platforms, that support publication of open data.

- 11.8 If third-party data is required for a JNCC-funded project, the project **must** make use of available open-licensed data in preference to non-open alternative data – provided the open-licensed data is fit for purpose and any risks associated with using the open-licensed data are no higher than those associated with the alternative.
- 11.9 JNCC **must** comply with cross-government policy on publication of open data on spending, expenses, contracts, tenders, pay structures, and organograms (the “transparency agenda”²⁶).

12 Communications

- 12.1 Any new publication of open data **should** be supported by external communications if the data has significant potential for public interest or the data significantly informs any policy or initiative that is or is likely to become a matter of public interest.
- 12.2 Staff responsible for external communications and public relations **must** be engaged prior to the release of any open data that is likely to create public controversy or reputational risk or prompt a significant level of public discussion.
- 12.3 JNCC’s Communications Team **should** be engaged prior to or shortly after the release of any JNCC open data that would benefit from promotion through social media or other communication channels.
- 12.4 Any new publication of open data that is likely to have high levels of re-use or significant impact, particularly

within biodiversity or conservation, **should** be supported by a fully developed communications plan.

- 12.5 Any external request for access to unreleased data held by JNCC **must** be handled in accordance with access to information law and JNCC’s internal guidance on access to information.
- 12.6 Data released in response to an access to information request **should** be accompanied by a statement of copyright and database right and terms for re-use, including an open licence if the data is eligible for re-use as open data.
- 12.7 Data that is released in response to an access to information request and is eligible for re-use as open data **should** also be made public through normal processes and channels for publication of open data.
- 12.8 Any request for re-use of data under the Re-use of Public Sector Information Regulations **must** be handled in accordance with those regulations. This policy **must** be applied to any release of data in response to such a request, except to the extent of any conflict between this policy and the regulations.
- 12.9 Open data **should** be supported at the organisational level by maintenance and open publication of an inventory of data assets. This inventory **should** be comprehensive and **should** include all data assets held, including those that are not eligible for external publication, subject to any redactions necessary for reasons of security, confidentiality or data protection. The inventory itself **should** be published as open data.
- 12.10 Open data **should** be supported at the organisational level by an easily discoverable online point of contact for submission of enquiries and

²⁶ <https://www.gov.uk/government/collections/how-to-publish-central-government-transparency-data>

comments about open data and requests for open release of data.

- 12.11 Any refusal made to an external request for release as open data of any data held by JNCC **must** be supported by a clear rationale.

13 Assessment and risk management

- 13.1 Any dataset considered for publication as open data by JNCC **must** be assessed before publication to determine its eligibility for publication and to identify any risks associated with public access to and re-use of the data. This assessment **must** be carried out in accordance with this policy and with any guidance on risk assessment of open data issued under this policy.

- 13.2 Guidance on risk assessment for open data publication will be issued by JNCC Digital and Data Solutions under this policy with approval from the Departmental Security Officer (DSO).

- 13.3 Risk assessment for open data publication has the following structure:

An initial risk assessment, carried out by the data manager, that **must** be approved by an Information Asset Owner (IAO). The initial risk assessment is sufficient to confirm eligibility of the dataset for publication as open data if the dataset conforms to a set of specified standard criteria.

A detailed risk assessment, carried out by or under the supervision of an IAO, that **must** be approved by an IAO or the DSO. The IAO **must** consult the DSO if the detailed risk assessment identifies any potential that publication of the dataset as open data will present a high risk.

If the detailed risk assessment determines that publication of the

dataset as open data presents any high risk, the assessment and any decision to publish the data, whether as open data or on more restrictive terms, **must** be approved by the DSO.

Any publication as open data of a dataset that contains material marked for handling as OFFICIAL-SENSITIVE **must** be approved by the DSO.

- 13.4 Publication of open data by JNCC **must** be supported by an internal record of the risk assessment undertaken including approval at the required level of responsibility.

- 13.5 Risk assessments **should not** normally be published with the open data. However documentation published with the open data **must** notify recipients of any risk identified by the assessment that could influence their decision to use the data.

- 13.6 Risk assessment for open data publication **should** be consistent with JNCC's approach to handling of requests for access to information. Any data that cannot be legally disclosed to the public in response to an access to information request **must not** be published as open data.

- 13.7 Sections 13.8-13.12 contain points about types of risk of particular concern to JNCC. Other types of risk are covered in JNCC's guidance on risk assessment.

13.8 Personal data

Risk assessment for open data publication **must** consider whether the dataset contains personal data or is derived from any source that contains personal data.

Risk assessment and processing of any dataset that contains personal data **must** follow JNCC's Data Protection Policy and **must** consider whether a Data Protection Impact Assessment (DPIA) is required.

All personal data in a dataset that will be published as open data **must** be either redacted, aggregated, or pseudonymised before publication so that individual data subjects cannot be or are very unlikely to be identified, unless there is a clear legal basis on which to disclose personal data in the published dataset.

Publication of any open data that discloses personal data **must** be approved by the DSO, subject to any more specific delegation of authority agreed by the DSO.

Any publication of open data that discloses personal data **must** be accompanied by documentation that notifies recipients of the disclosure and the legal basis on which it has been made.

13.9 Ecological harm and protection of the environment

A JNCC open data release **must not** involve any disclosure of information that would cause ecological harm or otherwise adversely affect the protection of the environment to which the information relates, unless JNCC has assessed the risk of harm and determined that the public interest in disclosing the information outweighs the potential for an adverse effect.

JNCC's approach to open data release of environmentally sensitive information **must** be consistent with JNCC's application of the exemption in regulation 12(5)(g) of the Environmental Information Regulations (EIR) when handling requests for access to information. In particular, ecological data and information listed as exempt from general release under EIR **must not** be published as open data.

This exemption usually applies to 'sensitive features' that could be put at risk if information about their location is made public. Sensitive features are "species, habitats or geological

formations which, due to factors such as rarity, fragility or attractiveness, are particularly vulnerable to harm caused by collecting, damage, disturbance or commercial exploitation."

The exemption will normally only be engaged if the disclosure is more likely than not to have an adverse effect. The public interest in favour of disclosure will be stronger if the adverse effect would not be particularly severe or would have a limited effect.

A dataset **must not** be withheld from publication as open data in its entirety simply because it contains some sensitive information. If a dataset contains information that cannot be published that information **must** be redacted before publication if the remaining information will be usable by itself.

13.10 Duty of confidence

Unless eligible for disclosure to the public in response to an access to information request, the following **must not** be published as open data:

any data or information classified or marked for handling as OFFICIAL-SENSITIVE, SECRET, TOP SECRET, or any equivalent marking or classification;

any data or information that would prejudice the confidentiality of proceedings if disclosed to the public;

any data or information received from another party where disclosure to the public would be a breach of confidence that is actionable; and

any data or information that would prejudice the commercial interests of JNCC or of another party if disclosed to the public.

13.11 Volunteered information (non-personal)

JNCC will sometimes publish as open data non-personal information provided on a voluntary basis by persons who were not under any legal obligation to supply the information to JNCC or to any other public authority.

For the purposes of this section a person may be an individual or a legal person such as a company.

By default JNCC's presumption is that volunteered information will be eligible for publication as open data unless:

disclosure of the information to the public would adversely affect the interests of the person who provided the information, and on that basis the information would be exempt from disclosure to the public in response to an access to information request; or

JNCC owes the person a duty of confidence with respect to the information and the person has not consented to disclosure of the information to the public.

13.12 Intellectual property

Data that is subject to third party intellectual property rights **must not** be published as open data, unless JNCC has a licence or other permission to publish that data under an open licence.

Data that has been created by any process that made use of third party intellectual property **must not** be published as open data, unless JNCC had a licence or permission to use that intellectual property for that purpose and has a legal right to publish the created data under an open licence.

A detailed risk assessment **must** be carried out and approved prior to the publication as open data of any dataset that contains data of unclear provenance or where it is unclear whether the dataset contains, or was created by a process that made use

of, suitably licensed third party intellectual property.

Any publication of open data that contains or may contain data of uncertain provenance or where rights to re-use third party intellectual property are in doubt **must** be accompanied by documentation that notifies recipients of those issues.

13.13 Risks that do not prevent publication as open data

The following risks or perceived risks **must not** by themselves prevent publication of open data:

- Concerns about data quality or quality assurance, provided the dataset meets minimum criteria for accessibility. It is generally understood that datasets published as open data have been created for a particular purpose and there is no warranty that they will be fit for wider re-use.
- Concerns about misinterpretation or misrepresentation of data by recipients, the media, or the public.
- Concerns about non-defamatory risk to the reputation of JNCC, other public bodies, or other parties.

13.14 Mitigation of risks

Any identified risk of publication of open data **must** be mitigated where practical if that mitigation will reduce the risk to a level that will enable publication of the dataset.

Risks or perceived risks that do not prevent publication of open data **should** also be mitigated where practical.

Standard forms of mitigation include redaction, aggregation, anonymisation or pseudonymisation, and information warnings or other explanations in documentation.

Annex A – Glossary

access control

For the purposes of this policy, access control is any mechanism used to regulate who, when, or how data can be accessed.

Access control includes any registration, identification, or authentication of recipients or users of data, data request forms, data sharing agreements, and any process for ordering or purchasing data.

access to information

For the purposes of this policy, access to information refers to the Freedom of Information Act, the Environmental Information Regulations, and any similar law that provides members of the public with a legal right to obtain information held by a public authority.

aggregation

the process of displaying data as totals, sometimes with suppression of small numbers

all rights reserved

'All rights reserved' is a formality used by a copyright holder to assert that they reserve for their own use all rights provided by copyright law. Under current UK and international law this formality is no longer legally necessary, but the phrase is still used commonly to indicate that a copyright holder does not permit re-use of their work.

By definition 'all rights reserved' cannot apply to open data or to any data that has been shared on terms that permit re-use.

anonymisation

the process of adapting data so that individuals cannot be or are unlikely to be identified

attribution and attribution statement

Attribution is acknowledgment of the source or authorship of a work.

In the context of data licensing an attribution statement is text that acknowledges the source of a dataset, usually in the form of a statement of copyright and database right including any rights of third parties that must be acknowledged.

Use of a specified attribution statement is a common condition for re-use of openly licensed data and other material.

citation

A citation is a reference to a published or unpublished source for the purpose of acknowledging in context the relevance of another work to the topic of discussion.

In the context of datasets, a citation is distinct from an attribution statement. A citation is likely to be in a prescribed style and focus on authorship of the dataset, particularly for purposes of scientific or scholarly credit, whereas an attribution statement will usually focus on acknowledgment of ownership.

citizen science

participation of the public in scientific work, including data gathering, usually on a voluntary basis and often in collaboration with professional scientists or scientific institutions

closed data

data that should not be shared and can only be accessed by its subject, owner or holder

copyright

an intellectual property right that gives the creator of an original work exclusive control over its use and distribution

data and dataset

For the purposes of this policy, data is facts and statistics collected together for reference or analysis, and a dataset is an organised collection of data.

This policy treats data as a singular noun because this is the standard usage within the open data community. However data is commonly treated as a plural in scientific writing and accordingly may be used as a plural in some documentation of JNCC datasets.

This policy does not make any nuanced distinction between data and information.

Data Protection Impact Assessment (PIA)

a tool that can help an organisation identify potential effects of a proposal or project on the privacy of individuals and to identify ways of ensuring compliance with data protection requirements

database right

an intellectual property right that enables the maker of a database to restrict extraction or re-use of substantial parts of the contents of the database, provided there has been substantial investment in obtaining, verifying or presenting the contents of the database

derived data

For the purposes of this policy, derived data is any data that has been created or produced through adaptation of data from one or more sources using calculation, aggregation, or some other type of transformation.

Derived data is a somewhat vague concept and the definition is subject to debate. Ordnance Survey has a particular view and rules around use of derived data but this approach is not necessarily applicable to data derived from other sources.

digital object identifier (DOI)

a persistent identifier assigned for the purpose of uniquely identifying a digital object

discoverability

the extent to which something is easy to find, and in particular the extent to which online content is easy to find via a search engine, on a website, or in an application

egress cost

the cost of transferring data out of a repository, usually in the context of storage hosted on the internet by a third party (cloud storage)

government data

data owned by national, regional, local, and city governments, international governmental bodies, and other types of institution in the wider public sector

human readable

A work is human readable if it is in a form that can be conveniently read by a human.

Data provided in a human readable form is not open data, unless that form is also machine workable or the data is separately available in a machine workable form.

information warning

An information warning is an advisory notice that communicates risks or caveats.

In the context of open data an information warning is documentation that highlights potential for misinterpretation or misuse of the data, and may include for example notes on data quality or purposes for which the data is unsuitable. An information warning will not impose any mandatory constraints on use of the data, but may draw attention to existing legal restrictions.

intellectual property (IP) and intellectual property rights (IPR)

Intellectual property is any creation of the mind that is recognised as an asset or property.

Intellectual property rights are legal rights granted to owners of IP that enable them to restrict use and distribution of the IP.

Copyright and database right are types of IPR of particular relevance in this policy. Other types of IPR include trademarks, patents, and design rights.

JNCC open data

any dataset eligible for publication as open data in which JNCC owns most of the intellectual property or has lead responsibility for creation of the dataset or for publication of the dataset

landing page

a human readable web page with no access barriers that provides information about a dataset or functions as a point of entry to a body of content on a particular theme

machine readable and machine workable

A work is machine readable if it is in a form that is readily processable by a computer.

A work is machine workable if it is in a form that is readily processable by a computer and where the individual elements of the work can be easily accessed and modified.

The distinction between these terms may be subjective and depend on the particulars of the work and its form. However machine workable is intended to denote a higher standard of accessibility.

marginal cost of reproduction

The marginal cost of reproduction (sometimes called the marginal cost of production or simply marginal cost) is the cost of producing one more unit of a good.

The marginal cost of reproducing a digital good such as a dataset will normally be very low and treated as nil for most practical purposes.

metadata

Metadata is data that describes or defines other data.

For the purposes of this policy metadata will typically conform to a recognised schema and is a type of documentation.

open and openness

Beyond their ordinary meaning, open and openness are used in a range of specific contexts (such as open data, open access, and open source) to refer to the general availability of material necessary for broad social and/or economic cooperation to occur in that context.

Open has a more specific definition in some contexts than in others. However open should not be confused with public; the public availability of material is not by itself sufficient to make that material open.

open access

Open access refers to research outputs (usually peer-reviewed) that are available online without restrictions on access and free of many restrictions on re-use.

Open access materials may include data, but data that is open access will not always conform to the definition of open data.

open API

an application programming interface that is publicly available with few restrictions and based on an open standard, and typically provides access to open-licensed data

open data

data that anyone is free to access, use, modify, and share for any purpose – subject, at most, to measures that preserve provenance and openness

Open Definition

The Open Definition ²⁷ is a key reference document, currently managed as a project of Open Knowledge International, ²⁸ that sets out principles that define "openness" in relation to data and content. The Open Definition has an open process for development and governance, and represents general consensus on the working definition of open data.

open by default

the principle that government data should be open unless there are legitimate reasons why the data cannot be released openly

open format

For the purposes of this policy, an open format is a file format, for storing digital data, which places no monetary or other restrictions on its use and can be fully processed with at least one software tool that is free of charge and generally available.

open licence and open-licensed

For the purposes of this policy, an open licence is a licence that conforms to the current version of the Open Definition.

Data or other material is open-licensed if it is re-usable under an open licence.

Note that licensing is only one part of the definition of open data, so open-licensed data will not always be open data.

open source

a decentralised model for software development that includes the principle that source code and documentation will be freely available to the public

open standard

For the purposes of this policy, JNCC will adopt the Free Software Foundation Europe's definition ²⁹ of an open standard as a format or protocol that is

1. subject to full public assessment and use without constraints in a manner equally available to all parties;
2. without any components or extensions that have dependencies on formats or protocols that do not meet the definition of an Open Standard themselves;
3. free from legal or technical clauses that limit its utilisation by any party or in any business model;
4. managed and further developed independently of any single vendor in a process open to the equal participation of competitors and third parties;
5. available in multiple complete implementations by competing vendors, or as a complete implementation equally available to all parties.

openwashing (or open-washing)

Openwashing is a term used to criticise the presentation of a product or practice as "open" when it does not conform to more widely accepted norms of openness.

In the context of open data, a data publisher risks accusations of openwashing if a dataset that does not conform to the minimum criteria for open data is presented as if it is open data.

²⁷ <http://opendefinition.org/>

²⁸ <https://okfn.org/>

²⁹ <http://fsfe.org/activities/os/def.html>

permissive

A licence is said to be permissive if it places few restrictions on how the material covered by the licence may be distributed or used.

All open licences are permissive but some are more permissive than others.

personal data

data or information relating to an identified or identifiable natural person

personally identifiable information (PII)

any information that can be used, on its own or with other information, to identify a specific individual

provenance

In general usage, provenance is the origin or history of something, such as a work of art.

In the specific usage in this policy, provenance is information the re-user needs about the source/s of a dataset, and sometimes about how the dataset was produced, so they can properly attribute the data and understand ownership of the data.

pseudonymisation

the process of attaching a coded reference or pseudonym to an individual in a dataset, with other identifying information removed

public domain

In intellectual property law, public domain refers to works in which intellectual property rights have expired, or have been forfeited or expressed waived, or are inapplicable.

In common usage, public domain refers to any material that is available to the public.

Please take care when using this term in discussions about open data, as there is potential for confusion between the two senses of meaning.

public task and public task information

Public task is a collective term for the core role and functions of a public body, including both statutory duties and any core role or function established through custom and practice.

Public task information is any data or information that a public sector body must produce, collect, or provide to fulfil its public task.

redaction

the process of removing or obscuring data or information for a legal or security purpose

reidentification or re-identification

the process of recovering personal data from anonymised data by using data matching or similar techniques

re-use and re-user

Re-use of material (including data and information) is use for any purpose other than that for which the material was produced. A re-user is a user who uses material for a purpose other than that for which the material was produced. A user of open data is usually a re-user, but open data licensing does not make any specific distinction between users and re-users.

schema

In the context of this policy, schema is mentioned in the context of metadata and refers to any standard markup vocabulary used to describe data in a structured manner that makes the data easier to discover.

shared data

data that is shared, either with particular individuals or organisations or for specific purposes or for public access, under terms and conditions that are not 'open'

signpost and signposting

As a verb, to signpost is to show the direction of something.

In the context of open data, signposting is the practice of making data more discoverable by highlighting its availability and where to find it. For example an article on a particular topic might include at the end a list of links to locations of related open datasets.

third party

In general usage, a third party is any person or group other than the two parties primarily involved in a situation.

In the specific usage in this policy, a third party is any person or organisation other than the primary copyright holder with intellectual property rights in the contents of a dataset.

transparency agenda and transparency dataset

The transparency agenda is a cross-government policy that mandates or encourages a range of administrative and financial datasets with the objective of promoting transparency and accountability of public services.

A transparency dataset is any dataset published in accordance with this agenda. Examples include datasets containing spending items or expenses, information on contracts and tenders, organograms, and other information on staff structures including grades and pay rates.

Annex B – Primer on intellectual property rights related to datasets

Licensing of open data depends on permissions granted by the organisation/s or individual/s that hold intellectual property rights (IPR) in the data.

Implementation of the Open Data Policy therefore requires a basic understanding of copyright and related rights that apply to datasets.

When a dataset is created as an original work it is subject to copyright.

Copyright is an intellectual property right that gives the creator of an original work exclusive control over its use and distribution. Copyright applies automatically provided that creation of the work required some degree of labour, skill or judgement.

Copyright is subject to some exemptions that allow use of publicly accessible information for limited purposes such as private study and non-commercial research.

Individual facts or observations are not subject to copyright, but a collection of facts or observations created as an original work is subject to copyright.

A new dataset may contain copyrighted material owned by third parties and used with their permission. Although the dataset owner will hold copyright over the dataset as a whole, their use of the data and their rights to allow others to use the data may be limited by terms and conditions from the third parties.

Database right is an intellectual property right that enables the maker of a database to restrict extraction or re-use of substantial parts of the contents of the database without their consent, provided there has been substantial investment in obtaining, verifying or presenting the contents of the database.

Database right is independent of copyright and does not require the maker of the database to have created any original work. This means the maker of a database may have an intellectual property right in a database without any copyright in the works collected in the database.

Database right exists in UK law and in the legal systems of other European Union states. However database right is not as widely recognised as copyright, and in some countries it does not exist as a legal concept. Database right does not exist in US law, for example.

Intellectual property rights are assets and may be transferred or sold by the rights owner to another party. Rights owners may also allow re-use of their IPR subject to terms and conditions of their choosing, including under an open licence.

Licensing of a dataset as open data, or under more restrictive terms, does not affect ownership of the data, or involve any assignment or transfer of intellectual property rights.

Publication of data or information on the web does not affect ownership of that material, or involve any assignment or transfer of intellectual property rights, and does not necessarily mean the publisher or owner of that material has granted members of the public any rights to re-use the material.

Annex C – Supporting information on the legal context to open data

Following are additional notes and links to more information about the legislation highlighted in section 4 of the Open Data Policy.

Re-use of Public Sector Information Regulations (RPSI or RoPSI)

The Re-use of Public Sector Information Regulations 2015
<http://www.legislation.gov.uk/ukxi/2015/1415/contents/made>

Directive 2013/37/EU on the re-use of public sector information
<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013L0037&from=FR>

Guide to RPSI (ICO)
<https://ico.org.uk/for-organisations/guide-to-rpsi/>

PSI Directive transposition and re-use regulations (National Archives)
<http://www.nationalarchives.gov.uk/information-management/re-using-public-sector-information/psi-directive-transposition-and-re-use-regulations/>

Notes:

- 1 RPSI does not require public bodies to make information available for re-use under an open licence. However it creates a formal process for requesting re-use of public sector information (PSI), with a complaints procedure and recourse to the courts. RPSI requires public bodies to justify restrictions they place on re-use of requested information.
- 2 National Archives is the main source of guidance on PSI including implementation of the legislation, and the Information Commissioner's Office (ICO) handles complaints about how public sector bodies have dealt with requests to re-use information.

Freedom of Information Act (FOI) and Environmental Information Regulations (EIR)

The Freedom of Information Act 2000
<https://www.legislation.gov.uk/ukpga/2000/36/contents>

The Freedom of Information (Scotland) Act 2002
<http://www.legislation.gov.uk/asp/2002/13/contents>

Guide to freedom of information (ICO)
<https://ico.org.uk/for-organisations/guide-to-freedom-of-information/>

Code of Practice (datasets) on the discharge of public authorities' functions under Part 1 of the Freedom of Information Act
<https://www.gov.uk/government/publications/secretary-of-states-code-of-practice-datasets-on-the-discharge-of-public-authorities-functions-under-part-1-of-the-freedom-of-information-act>

The Freedom of Information (Release of Datasets for Re-use) (Fees) Regulations 2013
<http://www.legislation.gov.uk/uksi/2013/1977/contents/made>

The Environmental Information Regulations 2004
<http://www.legislation.gov.uk/uksi/2004/3391/contents/made>

The Environmental Information (Scotland) Regulations 2004
<https://www.legislation.gov.uk/ssi/2004/520/contents/made>

Notes:

- 1 FOI and EIR cover requests for access to information, and in general do not cover licensing of information for re-use. Section 11A of the Freedom of Information Act 2002, which took effect in 2013, contains specific provisions on release of datasets for re-use. However 'dataset' has a narrower definition in this Act than it does in common usage, and in practice section 11A has limited effect because it excludes copyrighted works to which RPSI applies.
- 2 Section 11 of the Freedom of Information Act 2002 states that a public authority must provide requested information in an "electronic form which is capable of re-use" (if requested and as far as reasonably practicable). This requirement was introduced to address a common complaint that public authorities were releasing data in formats that made the data difficult to analysis. There is no similar language in EIR; however EIR creates a duty to make environmental information available proactively using easily accessible electronic means, and failure to provide requested information in a re-usable format may conflict with that duty.

Data Protection

The Data Protection Act 2018
<http://www.legislation.gov.uk/ukpga/2018/12/enacted>

The General Data Protection Regulation
<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>

Guide to data protection (ICO)
<https://ico.org.uk/for-organisations/guide-to-data-protection/>

Guide to the General Data Protection Regulation (GDPR) (ICO)
<https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr>

2018 reform of EU data protection rules (EU)
https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

Notes:

- 1 The GDPR applies in the UK from 25 May 2018. The Data Protection Act 2018 is new law that makes some provisions about how the GDPR applies in the UK, and makes other provisions about areas of data protection that are outside the scope of the GDPR. The Data Protection Act 1998 was repealed when the GDPR and the new Data Protection Act took effect.

INSPIRE

INSPIRE Directive (2007/2/EC)

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:en:PDF>

INSPIRE Knowledge Base (EU)

<https://inspire.ec.europa.eu/>

The INSPIRE Directive 2007 (Data.gov.uk)

<https://data.gov.uk/location/inspire>

Annex D – Assessment tools and marks of quality

The 5-star deployment scheme for Open Data

The 5-star deployment scheme for Open Data, ³⁰ also known as the “five stars of openness”, is a scale for describing presentation of open data and linked open data.

The scheme is derived from a post ³¹ written by Tim Berners-Lee in 2006 and last updated in 2010. Implementation of the scheme is open. Stars are awarded to datasets on the following basis:

☆	Available on the web (whatever format) but with an open licence, to be Open Data
☆☆	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
☆☆☆	as (2) plus non-proprietary format (e.g. CSV instead of excel)
☆☆☆☆	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
☆☆☆☆☆	All the above, plus: Link your data to other people’s data to provide context

Notes:

- 1 The scale is cumulative. For example the third star is only awarded if the dataset also meets the criteria for the first two stars.
- 2 Datasets with only one or two stars are not open data within the generally accepted definition. Datasets with three stars are open data. The fourth and fifth stars are refinements based on open standards and linking the dataset to other data.
- 3 There is no expectation that every open dataset should have more than three stars. Some datasets are not suitable for conversion to linked data and/or the publisher may not be able to justify the resources required to make all datasets eligible for more than three stars.
- 4 The examples of file formats used in the original version of the scheme are out of date, as Excel formats are now open standards and supported by free software. Subject to minimum standards of accessibility in presentation of the data, JNCC will treat data published in XLS and XLSX formats as eligible for the third star.
- 5 JNCC will implement this scheme based on the criteria under which JNCC open data is published. The results for individual datasets may not coincide with the previous implementation of the scheme on Data.gov.uk ³² (DGU), as that was based on automated detection of file formats by the DGU platform.

³⁰ <http://5stardata.info/en/>

³¹ <https://www.w3.org/DesignIssues/LinkedData.html>

³² http://webarchive.nationalarchives.gov.uk/20160811144258/http://guidance.data.gov.uk/five_stars_of_openness.html

Open Data Certificates

Open Data Certificates³³ are a tool devised by the Open Data Institute for self-certification of individual open datasets based on different levels of good practice.

This system is designed to provide data users with an assurance of the extent to which an open dataset is dependable and adheres to recognised standards.

JNCC will endeavour to achieve Gold-level certification for business-critical datasets.

The Open Data Pathway

The Open Data Pathway is a tool developed by the Open Data Institute with Defra to enable organisations to assess their open data maturity.

JNCC will use the Pathway to periodically assess the organisation's progress on open data policy and the business change required to support it. Outputs from these assessments will be reported to JNCC's Executive Management Board so that EMB members can prioritise areas of improvement amongst the five themes in the maturity model that underpins the tool.

The five themes in the Open Data Maturity Model³⁴ are:

- 1 Data management processes: identifies the key business processes that underpin data management and publication including quality control, publication workflows, and adoption of technical standards.
- 2 Knowledge & skills: highlights the steps required to create a culture of open data within an organisation by identifying the knowledge sharing, training and learning required to embed an understanding of the benefits of open data.
- 3 Customer support & engagement: addresses the need for an organisation to engage with both their data sources and their data re-users to provide sufficient support and feedback to make open data successful.
- 4 Investment & financial performance: covers the need for organisations to have insight into the value of their datasets and the appropriate budgetary and financial oversight required to support their publication. In terms of data consumption, organisations will need to understand the costs and value associated with their re-use of third-party datasets.
- 5 Strategic oversight: highlights the need for an organisation to have a clear strategy around data sharing and re-use, and an identified leadership with responsibility and capacity to deliver that strategy.

³³ <https://certificates.theodi.org/en/>

³⁴ <https://theodi.org/article/open-data-maturity-model/>

Annex E – Supporting notes on formats and accessibility

The most accessible file format for publication of an open dataset will depend on the type of data and the particulars of the dataset, and may also depend on the context in which the data is published.

Section 7 of the Open Data Policy provides criteria for recognising and selecting a file format but does not specify individual file formats. However following are some guidelines:

- 1 Large tabular datasets should ideally be published in CSV or TSV format. Structured TXT formats are also fine as long as the schema is clearly documented.
- 2 Spatial datasets should be published in shapefile (SHP), GML, KML, or GeoJSON. Other formats may also be suitable. There are currently doubts around the accessibility of geodatabases (GDB).
- 3 Microsoft's main Excel formats XLS and XLSX (Office Open XML) are supported by alternative free software such as OpenOffice, so are generally fine for publication of open data – provided the data is suitably presented. Alternative software may struggle with VBA (macros), annotation in cells (comments), colour-coding used to represent information, and data in pivot tables. Plainer presentations are less likely to present problems.
- 4 Data in a PDF is rarely accessible enough to be open data – even if the contents of the PDF are covered by an open licence. But there's nothing wrong with presenting data and other information in PDFs for audiences who just want a readable version.
- 5 If there isn't an obvious choice of file format that will keep everyone happy, think about publishing open data in multiple formats; resources permitting.

Annex F – Supporting notes on attribution statements

The Open Government Licence and most other open licences require the licensee to attribute the source of the information covered by the licence if they re-use the information in a product or application. The publisher of an open dataset should specify the attribution statement they require.

This is a standard attribution statement for a JNCC dataset:

Contains Joint Nature Conservation Committee data © copyright and database right 2018

If the dataset contains third-party data (included with permission) the attribution statement must also include any attribution required by the third party. For example:

Contains Joint Nature Conservation Committee data © copyright and database right 2018
Contains Scottish Natural Heritage data © copyright and database right 2018

It's okay to abbreviate "Joint Nature Conservation Committee" as "JNCC", and to abbreviate the names of other organisations provided they have no objection and the meaning remains clear. For example:

Contains JNCC data © copyright and database right 2018
Contains OS data © Crown copyright and database right 2018

It is best to make an attribution statement as short as possible. Long statements can create difficulty for users, in particular when displayed on maps or in reports.

Government departments and some other UK public sector bodies are Crown bodies ³⁵ and covered by "Crown copyright". However JNCC is not. JNCC is covered by normal copyright.

If an attribution statement applies to information or other content that does not include structured data there is no need to mention "database right".

Some copyright statements and attribution statements include the phrase "all rights reserved" to indicate the copyright holder does not allow re-use. Attribution statements for open data should never contain "all rights reserved".

There is some flexibility in how attribution statements are written. The important thing is to clearly show the source/s of the information and acknowledge the rights of any third parties that have intellectual property rights in the information. However an error in writing an attribution statement will not change the legal effect of any intellectual property rights.

³⁵ <https://bit.ly/2uyB85T>

Acknowledgements

Some definitions used in this Open Data Policy and in Annex A – Glossary are copied or adapted from the following sources:

The Open Definition (OKFN)
<http://opendefinition.org/>

The Open Data Institute's 'data spectrum' model
<https://theodi.org/about-the-odi/the-data-spectrum/>

The Glossary of Public Sector Information and Open Data Terminology (Data.gov.uk)
<http://webarchive.nationalarchives.gov.uk/20161108184844/https://data.gov.uk/glossary>

The definition of 'sensitive features' in section 13.9.3 of this policy is taken from the following source:

Environmental Information Regulations Guidance Note No 1 (Countryside Agencies' Open Information Network)
The 'Environmental Exception' and access to information on sensitive features
<https://nbn.org.uk/wp-content/uploads/2016/03/EIR-Guidance-on-the-Environmental-Exception-1.pdf>

The definition of 'open standard' in Annex A – Glossary follows the Free Software Foundation Europe's definition: <http://fsfe.org/activities/os/def.html>.