



# Open Data Technical Framework

Developed in collaboration with the Public Bodies Working Group on Open Data

Government Reform Unit

Department of Public Expenditure and Reform

June 2015



## Contents

Acknowledgement .....	3
Background .....	4
Need for an Open Data Technical Framework.....	4
Publishing Open Data.....	5
Open Data Licence .....	7
Recommended Formats for Open Data.....	8
Recommended Metadata Schema for Open Data.....	10
Dcat:Dataset .....	11
gmd:EX_GeographicBoundingBox .....	11
Recommended Standards for Open Data.....	13
Unique Resource Identifiers .....	16
Annex 1: Open Data Glossary .....	17
Annex 2: Open Data Checklist.....	20

## Acknowledgement

This document was prepared by the Public Bodies Working Group (PBWG) on Open Data. The Department of Public Expenditure and Reform wishes to acknowledge the important contribution of and to thank the PBWG members for their commitment and hard work over the last 6 months in developing this framework which underpins the publication of datasets on the portal ([data.gov.ie](http://data.gov.ie)) in line with agreed standards thus facilitating re-use and interoperability:

Eoin McCuirc, Central Statistics Office

Dominic Byrne, Fingal County Council

Gareth John, Department of Arts, Heritage and the Gaeltacht

Liam Stewart, Office of Public Works

Niall Hayden, National Roads Authority

John Nott, National Transport Authority

Sandra Collins, Digital Repository of Ireland

Rebecca Grant, Digital Repository of Ireland

Tracey P. Lauriault, Programmable Cities, NUI Maynooth

Pat Mulhall, Office of the Revenue Commissioners

Keith Walsh, Office of the Revenue Commissioners

Eoin O'Grady, Marine Institute

Rob Ovington, Department of Environment, Community and Local Government

Brian Costello, Central Statistics Office

Adam Leadbetter, Marine Institute

Hugh Mangan, Ordnance Survey Ireland

Ken Noble, Ordnance Survey Ireland

Vincent Hussey, Office of Public Works

Martin Troy, Department of Public Expenditure and Reform

Evelyn O'Connor, Department of Public Expenditure and Reform (Chair)

With technical support and advice from Deirdre Lee, Derilinx

## Background

Open Data is recognised as a key element of the Public Service Reform agenda and improved data management is an important element of a wide variety of key policy documents and action plans. Ireland has also committed to meeting the challenges set under the G8 Open Data Charter.



## Need for an Open Data Technical Framework

A key priority under the Minister for Public Expenditure and Reform, Mr. Brendan Howlin T.D.'s Open Data Initiative is the development and expansion of the National Open Data Portal, <http://data.gov.ie>. The objective of the Open Data Portal is to publish government data in a way that will make it more discoverable, accessible, interoperable and reusable. The Portal has been updated to support the publication of high-value datasets to meet demand and contribute to the achievement of real economic, social and democratic benefits for citizens, business and the Public Sector.

This document sets out the proposed technical framework that supports the ongoing implementation of the Open Data Initiative and ensures that publication of datasets on the Open Data Portal, [data.gov.ie](http://data.gov.ie), is done in a consistent, persistent and truly open way. This is a living document that will be expanded upon as technologies and practices evolve.

This Technical Framework comprises five key components:

1. Open Data Licence
2. Recommended Formats
3. Metadata Schema
4. Recommended Standards
5. Recommended Unique Resource identifiers

## Publishing Open Data

The Open Data Technical Framework sets out a planned and structured approach to the publication of datasets as Open Data. Public Bodies, when considering publication of Open Data should take into account the value, potential for re-use, and contribution datasets can make to delivering better outcomes for citizens, business, and other public servants and to help improve evidence-based decision making by public bodies.

Decisions on publication of Open Data will ultimately be a matter for individual public bodies, following Data Audits which will be conducted in all public bodies over time.

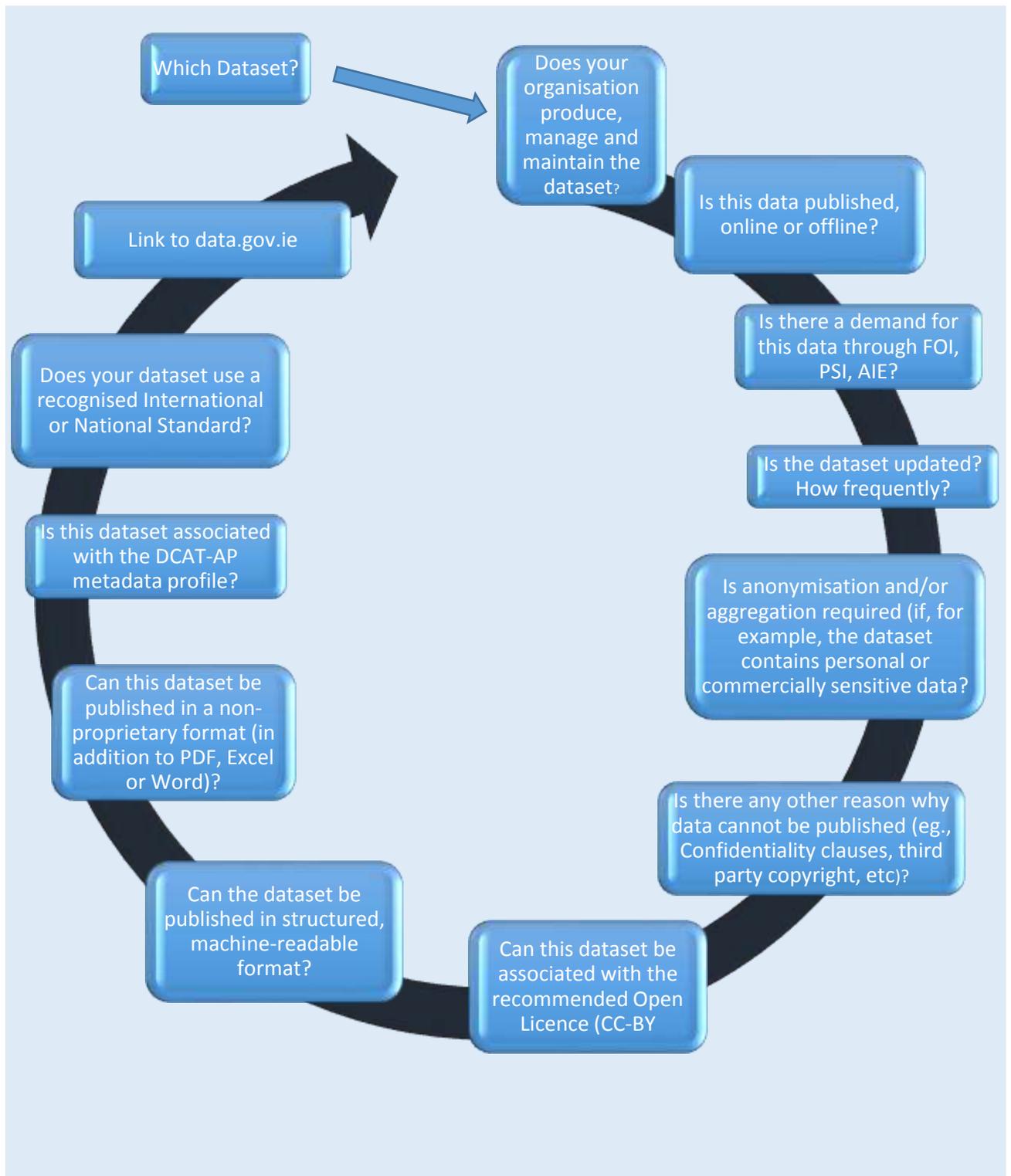
Data Audits are important in that they form the basis for a planned and structured approach to be taken to the publication of data as Open Data; taking into account the value, potential for re-use and contribution it can make to help achieve Public Service reform and national economic objectives. More generally, auditing of datasets should be seen as part of an organisation's information management strategy.

The output of audits will facilitate publication of datasets on our national portal, <http://data.gov.ie>. Audits will also promote the effective management and use / sharing of information in public bodies and support the implementation of the Public Service ICT strategy. Audits will enable identification of:

- The extent and range of datasets that exist and are managed and maintained by a public body
- The ranking of datasets in terms of their importance to the delivery of Departmental objectives and the perceived gaps in useful data that might help the delivery of these objectives
- The potential for sharing datasets within the Department and the wider public sector
- The potential for publication and making available for re-use – obligations in this regard will increase once the amended PSI Directive has been transposed and it would also be expected that greater publication and access to open data would reduce reliance on access to information under FOI
- To build on the recommendations of the 2012 IMF Board paper by promoting the placing of greater emphasis on the quality of reported data/information.

A high level decision process map setting out the key issues to be considered is included as Table 1 below. An Open Data publication checklist is included at Annex 2.

Table 1 – Decision Process Map



## Open Data Licence

For a dataset to be considered as Open Data, it must be published under an Open Licence. The European Commission, as part of its ongoing work in relation to the Revision of the PSI Directive, has issued guidelines on recommended licences and datasets. These guidelines encourage “the use of open licences, which should eventually become common practice across the Union”.

Following a public consultation on options for Ireland’s Open Data Licence, 14 responses were received. There was broad support for the use of the Creative Commons Attribution (CC-BY) licence from respondents. This licence lets others distribute, remix, tweak, and build upon data, even commercially, as long as users credit the original publisher for the original creation. CC-BY<sup>1</sup> is recommended for maximum dissemination and use of licensed materials. The proposed licence statement and recommended disclaimer statements should be used under the Open Data Initiative.

### **Recommendations:**

All data and metadata linked to data.gov.ie will be associated with the Creative Commons Attribution (CC-BY) Licence, at a minimum. Public bodies may waive copyright and associate datasets with CCO, if that is considered appropriate. The licence should be clearly identified in the metadata.

Only datasets associated with the recommended Open Data Licence may be included on data.gov.ie. However, datasets clearly associated with another licence, such as the PSI Licence, may be linked to the Open Data portal provided a commitment is made to using the Open Standard licence within a clearly defined timeframe.

### **Licence Statement**

Under the CC-BY Licence, users must acknowledge the source of the Information in their product or application by including or linking to this attribution statement: *“Contains Irish Government Data licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence”*.

### **Multiple Attributions**

If using data from several Information Providers and listing multiple attributions is not practical in a product or application, users may include a URI or hyperlink to a resource that contains the required attribution statements.

### **Disclaimer**

All data linked to the Open data portal is published “as is”. The Information is licensed 'as is' and the Information Provider and/or Licensor excludes all representations, warranties, obligations and liabilities in relation to the Information to the maximum extent permitted by law.

The Information Provider and/or Licensor are not liable for any errors or omissions in the Information and shall not be liable for any loss, injury or damage of any kind caused by its use. The Information Provider does not guarantee the continued supply of the Information.

### **Exemptions**

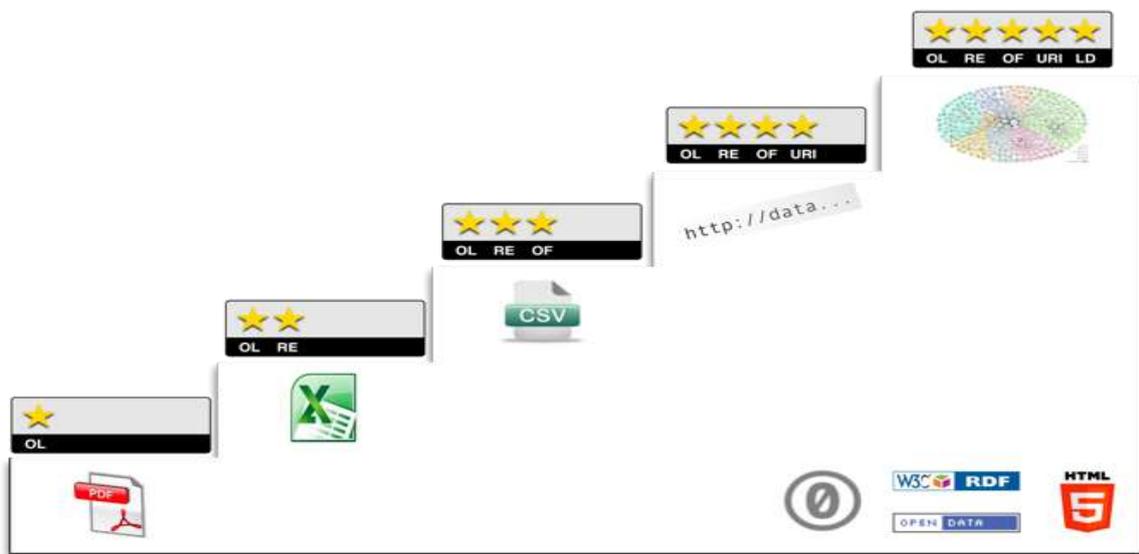
This licence does not cover personal data in the information. Nor does it cover third party rights (including, but not limited to, patents, copyright, database rights or trademarks).

---

<sup>1</sup> <http://creativecommons.org/licenses/>

## Recommended Formats for Open Data

While data published in any format can be considered Open Data if associated with an Open Licence, the type of data format used can have significant implications for the usability of the data. Under the Open Data Initiative, public bodies should publish their data in the most open way possible. One way to measure the openness of the formats used is through the 5-star deployment scheme for Open Data<sup>2</sup>. The greater the number of stars, the more reusable the data.



5-Star Open Data Scheme<sup>3</sup>

For inclusion in the Open Data Portal, public bodies must publish data at a minimum of 3 Star Open Data, such as CSV, JSON or XML. However it is encouraged to publish datasets in multiple formats, for example, 1 Star (e.g. PDF), 2 Star (e.g. Microsoft Excel) in addition to the required 3 Star (e.g. CSV).

### Recommendations

All datasets on <http://data.gov.ie> should be available in at least one of the following formats:

#### General

CSV  
JSON  
XML  
ODF  
RDF

#### Geospatial

GeoJSON  
GML  
KML  
WKT  
LAS  
IFC  
Shapefile  
ASCII Grid

#### Domain Specific

NetCDF  
Datex II  
GTFS  
JSON STAT

This list is subject to review and updating as new formats are developed due to technological developments.

<sup>2</sup> <http://www.w3.org/DesignIssues/LinkedData.html>

<sup>3</sup> <http://5stardata.info/>

If a public body intends to change a publication format, it should give prior notice to users stating the date from which the new format will be introduced. Sufficient time (three months suggested) should be given to allow users to make any necessary arrangements to ensure that they are not adversely affected by the change.

Datasets may be published in multiple formats

Datasets not yet available in one of the recommended open formats should have a clear timeframe when it will be available in an open format.

The publication of data in open formats should be built into data publication processes of all public bodies, and as part of information management more generally.

## Recommended Metadata Schema for Open Data

In order to help realise the benefits of Open Data, public bodies should make their data more searchable and usable. To achieve this, public bodies should provide precise descriptors about their datasets to help in the identification, location and retrieval of online resources by data-users.

These descriptors are commonly known as “metadata”.

Metadata is the summary information describing the data, including the availability, nature and constituents of the data. It provides context about the data that helps users understand their meaning, such as:

- What is the dataset called?
- What is the subject matter?
- Where can I locate the dataset?
- When was it produced and last updated?
- From what sources was the information compiled?
- Are there any restrictions on their use?

The Open Data Initiative requires a consistent approach to the publication of Open Data to ensure interoperability between datasets published by public bodies, at both national and international levels.

Accordingly, this Technical Framework recommends the adoption of a standardised Metadata Schema by public bodies, namely the **DCAT Application Profile (DCAT-AP)**, which is being used in a number of European Open Data portals; with appropriate geospatial extensions. An extracted Reference Guide to DCAT-AP is at Table 2 below and the complete specification is available at: [https://joinup.ec.europa.eu/asset/dcat\\_application\\_profile/home](https://joinup.ec.europa.eu/asset/dcat_application_profile/home)

One aspect of DCAT-AP that is lacking is geospatial metadata coverage. The EU DCAT-AP Working Group has identified the need to describe geospatial datasets, data series, and services. As a result, that Group is working on GeoDCAT-AP<sup>4</sup>, an extension of DCAT-AP. For the purpose of data.gov.ie, the geospatial metadata properties defined in Table 3 will be included.

**Table 2: DCAT-AP Reference Guide to Catalogue, Dataset and Distribution Classes<sup>5</sup>**

Class	Class URI	Mandatory properties	Recommended properties	Optional properties
<b>Catalogue</b>	dcat:Catalog	dcat:dataset dct:description dct:publisher dct:title	dct:issued dct:language dct:license dct:modified dcat:themeTaxonomy foaf:homepage	dcat:record dct:rights dct:spatial dct:isPartOf dct:hasPart
<b>Dataset</b>	dcat:Dataset	dct:description dct:title	dcat:contactPoint dcat:distribution dcat:keyword	dct:conformsTo dct:creator dct:accrualPeriodicity

<sup>4</sup> <https://joinup.ec.europa.eu/node/139283/>

<sup>5</sup> [https://joinup.ec.europa.eu/asset/dcat\\_application\\_profile/description](https://joinup.ec.europa.eu/asset/dcat_application_profile/description)

			dcat:theme dcat:publisher	dct:identifier dcat:landingPage dct:language adms:identifier dct:issued dct:spatial dct:temporal dct:modified owl:versionInfo adms:versionNotes dct:provenance dct:source dct:accessRights dct:hasVersion dct:isVersionOf dct:relation foaf:page adms:sample dct:conformsTo
<b>Distribution</b>	dcat:Distribution	dcat:accessURL	dct:description dct:format dct:license	dcat:byteSize spdx:checksum dcat:downloadURL dcat:mediaType, dct:issued dct:rights adms:status dct:title dct:modified dct:conformsTo foaf:page dct:language

**Table 3: Geospatial metadata elements to be included in data.gov.ie**

Property	URI	Domain	Range	Usage note	Card.
<b>Geographic Bounding Box</b>	gmd:EX_GeographicBoundingBox	Dcat:Dataset	gmd:EX_GeographicBoundingBox	<a href="http://www.datypic.com/sc/niem20/t-gmd_EX_GeographicBoundingBox_Type.html">http://www.datypic.com/sc/niem20/t-gmd_EX_GeographicBoundingBox_Type.html</a>	0..1
<b>Temporal Extent</b>	dct:temporal	Dcat:Dataset	dct:PeriodOfTime	This property refers to a temporal period that the Dataset covers.	0..n
<b>Lineage</b>	dct:provenance	Dcat:Dataset	dct:ProvenanceStatement	This property contains a statement about the lineage of a Dataset.	0..n
<b>Spatial Reference</b>	gmd:MD_ReferenceSystem	Dcat:Dataset	gmd:MD_ReferenceSystem	<a href="http://www.datypic.com/sc/niem20/e-">http://www.datypic.com/sc/niem20/e-</a>	0..1

<b>ce System</b>			As defined in Table	gmd_MD_ReferenceSystem.html	
<b>Spatial Resolution</b>	gmd:MD_Resolution	Dcat:Dataset	gmd:MD_Resolution	http://www.datypic.com/sc/niem20/e-gmd_MD_Resolution.html	0..1
<b>Conformance</b>	dct:conformsTo	Dcat:Dataset	dct:Standard	This property refers to an implementing rule or other specification.	0..n
<b>Complete Metadata Profile</b>		Dcat:Dataset			

**Table 4: Spatial Reference Systems**

Name	Acronym	EPSG Code	URI
<b>Irish Transverse Mercator</b>	ITM	2157	http://www.opengis.net/def/crs/EPSSG/0/2157
<b>Irish Grid</b>		29902	http://www.opengis.net/def/crs/EPSSG/0/29902
<b>European Terrestrial Reference System 1989</b>	ETR89	4258	http://www.opengis.net/def/crs/EPSSG/0/4258
<b>Ireland 1975 Mapping Adjustment</b>		(1953/1956?)	
<b>World Geodetic System 1984</b>	WGS-84	4326	http://www.opengis.net/def/crs/EPSSG/0/4326

## Recommendations:

All Open Data must be associated by standardised metadata.

All metadata must be accompanied by the Open Licence.

**DCAT AP** will be adopted as the Open Data Initiative's Metadata Schema, with appropriate geospatial values outlined.

All datasets on <http://data.gov.ie> will be accompanied by metadata compliant to DCAT-AP (with the Geo extension, if appropriate to the dataset).

This Metadata Schema includes three categories of metadata, as set out in Table 3 above:

- Mandatory
- Recommended
- Optional

## Recommended Standards for Open Data

Data standards, also referred to as data models or data vocabularies, ensure a common understanding of data content and what it describes to data users; and help facilitate the smooth exchange of data. Standards ensure that data is published in a permanent, persistent and consistent way.

Data standards help give a common meaning to data. This is especially important when data is being used by a third-party, being integrated from different sources, or when data is being shared across public bodies. Data standards not only define the meaning of certain concepts, but also how concepts relate to each other, which facilitates data interoperability.

When publishing Open Data, international standards defined by reputable standards organisations, such as ISO, the European Commission, W3C, IETF, OGC and OASIS should be used if possible. If international standards are unavailable or unsuitable, use national standards. For specific topics such as geospatial, statistics, or health, use national standards as defined by the responsible organisation (OSI, CSO, HIQA, etc.).

The Public Bodies Working Group (PBWG) reviewed the commonly used data standards by Irish Public Bodies. These are defined in Table 5. This is not an exhaustive list and is designed to be a go-to point for data publishers. The list will be updated with new standards as they are adopted in general practice.

**Table 5: Recommended Data Standards for data.gov.ie**

Short Title	Title	Domain	Standardisation Body	URL
<b>AR-DRG</b>	Australian Refined Diagnosis Related Group	Health	Australian Government	<a href="http://www.aihw.gov.au/hospitals-data/ar-drg-data-cubes/">http://www.aihw.gov.au/hospitals-data/ar-drg-data-cubes/</a>
<b>ATC/DDD</b>	The Anatomical Therapeutic Chemical Classification System with Defined Daily Doses	Chemical	WHO	<a href="http://www.who.int/classifications/atcddd/en/">http://www.who.int/classifications/atcddd/en/</a>
<b>COICOP</b>	Classification of Individual Consumption According to Purpose	Consumption	UN Statistics Division	<a href="http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5">http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5</a>
<b>CSO Standard Classifications</b>	CSO Standard Classifications	Multiple	CSO	<a href="http://www.cso.ie/en/surveysandmethodology/classifications/standardclassifications/">http://www.cso.ie/en/surveysandmethodology/classifications/standardclassifications/</a>
<b>CSO Standards</b>	CSO Standard Classifications	Statistics	CSO	<a href="http://www.cso.ie/en/surveysandmethodology/classifications/standardclassifications/">http://www.cso.ie/en/surveysandmethodology/classifications/standardclassifications/</a>
<b>DataCube</b>	Data Cube Vocabulary	Statistical	W3C	<a href="http://www.w3.org/TR/vocab-data-cube/">http://www.w3.org/TR/vocab-data-cube/</a>
<b>DCAT</b>	Data Catalog Vocabulary	Metadata	W3C	<a href="http://www.w3.org/TR/vocab-dcat/">http://www.w3.org/TR/vocab-dcat/</a>
<b>DCMI</b>	Dublin Core Metadata Initiative	Metadata	Dublin Core	<a href="http://dublincore.org/documents/dcmi-terms/">http://dublincore.org/documents/dcmi-terms/</a>
<b>Disadvantage Index</b>	Disadvantage index		ERC	?

<b>EUCAN</b>	Common Cancers	Cancer	WHO	<a href="http://eco.iarc.fr/eucan/Default.aspx">http://eco.iarc.fr/eucan/Default.aspx</a>
<b>IANA</b>	IANA Media Types	Media/File Types	Internet Assigned Numbers Authority	<a href="http://www.iana.org/assignments/media-types/media-types.xhtml">http://www.iana.org/assignments/media-types/media-types.xhtml</a>
<b>IATI</b>	International Aid Transparency Initiative	Transparency	IATI	<a href="http://iatistandard.org/">http://iatistandard.org/</a>
<b>ICCS</b>	Irish Crime Classification System	Crime	CSO	<a href="http://www.cso.ie/en/media/csoie/releasespublications/documents/crimejustice/current/crimeclassification.pdf">http://www.cso.ie/en/media/csoie/releasespublications/documents/crimejustice/current/crimeclassification.pdf</a>
<b>ICD</b>	International Classification of Diseases	Health	WHO	<a href="http://www.who.int/classifications/icd/en/">http://www.who.int/classifications/icd/en/</a>
<b>ISO 19100</b>	19100 Geographic Information standard series developed by the International Organization for Standardization (ISO)	Geospatial	ISO/OGC	<a href="http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=54904&amp;published=on&amp;includesc=true">http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=54904&amp;published=on&amp;includesc=true</a>
<b>INSPIRE</b>	Infrastructure for Spatial Information in the European Community	Spatial / Environmental	EC	<a href="http://inspire.ec.europa.eu/">http://inspire.ec.europa.eu/</a>
<b>ISO 3166-2:IE</b>	Country codes and subdivisions		ISO	<a href="http://www.iso.org/iso/iso_3166-2_newsletter_ii-3_2011-12-13.pdf">http://www.iso.org/iso/iso_3166-2_newsletter_ii-3_2011-12-13.pdf</a>
<b>ISO 639</b>	Language codes	Language	ISO	<a href="http://www.iso.org/iso/home/standards/language_codes.htm">http://www.iso.org/iso/home/standards/language_codes.htm</a>
<b>ISO 8601</b>	Date and time format	Date/Time	ISO	<a href="http://www.iso.org/iso/home/standards/iso8601.htm">http://www.iso.org/iso/home/standards/iso8601.htm</a>
<b>ISO 4217</b>	Currency codes	Multitple	ISO	<a href="http://www.iso.org/iso/home/standards/currency_codes.htm">http://www.iso.org/iso/home/standards/currency_codes.htm</a>
<b>MDC</b>	Major Diagnostic category	Health	Utah Department of Health	<a href="http://health.utah.gov/opha/IBISHelp/codes/MDC.htm">http://health.utah.gov/opha/IBISHelp/codes/MDC.htm</a>
<b>NACE Rev.2</b>	NACE Rev.2	Metadata	Eurostat	<a href="http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&amp;StrNom=NACE_REV2&amp;StrLanguageCode=EN&amp;IntPcKey=&amp;StrLayoutCode=HIERARCHIC&amp;CFID=1110191&amp;CFTOKEN=3ca0f6dad71d377-1F2DE4F0-F7BF-BCAE-31C18C386EA88F92&amp;jsessionId=f900daad75c14b465532m">http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&amp;StrNom=NACE_REV2&amp;StrLanguageCode=EN&amp;IntPcKey=&amp;StrLayoutCode=HIERARCHIC&amp;CFID=1110191&amp;CFTOKEN=3ca0f6dad71d377-1F2DE4F0-F7BF-BCAE-31C18C386EA88F92&amp;jsessionId=f900daad75c14b465532m</a>
<b>NUTS</b>	Nomenclature of territorial units for statistics		EC	<a href="http://ec.europa.eu/eurostat/web/nuts/overview">http://ec.europa.eu/eurostat/web/nuts/overview</a>
<b>SDMX</b>	Statistical Data and Metadata eXchange	Statistical	SDMX	<a href="http://sdmx.org/">http://sdmx.org/</a>
<b>Total poverty index</b>	Total poverty index		ERC	?
<b>XBRL</b>	eXtensible Business Reporting Language	Business	XBRL	<a href="http://www.xbrl.org/">http://www.xbrl.org/</a>
	The 26 geographic counties, except that Tipperary NR and Tipperary SR are distinguished		CSO	?

The 26 geographical counties	CSO	?
The 32 geographic counties of Ireland and Northern Ireland	CSO	?
The 34 administrative counties, except that Tipperary NR and Tipperary SR are combined	CSO	?
The 34 administrative counties	CSO	?

## Recommendations

Use Table 5 as a reference of data standards commonly used for Open Data in Ireland.

When publishing Open Data, public bodies should first try to reuse international standards defined by reputable standards organisations, such as ISO, the European Commission, W3C, IETF, OGC and OASIS.

If international standards are unavailable or unsuitable, use national standards. For specific topics such as geospatial, statistics, or health, promote national standards defined by the responsible organisation (OSI, CSO, HIQA, etc.).

## Unique Resource Identifiers

The Technical Framework sets out steps that will allow public bodies to achieve a minimum 3 Star Open Data. In the longer term, however, the intention is to progress to greater levels of linked data (4 and 5 Star).

The use of Universal Resource Identifiers (URIs) is an important element of this longer term approach to Open Data and the Public Bodies Working Group will be tasked with developing a URI Strategy and agreeing a URI pattern for use under the Open Data Initiative, using international experience and best practice.

The ongoing development of Open Data and the desire to increase its interoperability have led to an increased reliance on URIs as identifiers for a wide variety of concepts; everything from languages to buildings, public bodies to currencies. URIs are valuable in that they can help distinguish data resources and facilitate unique data identification, comparison and linking. URIs can be used to identify anything from places and people to things and concepts.

It is intended that the outcome of the work of the PBWG will be persistent and scalable URI patterns that will continue to be used even when public bodies change and applications using URIs expand.

### Examples of URI patterns

- General Pattern: `http://{domain}/{type}/{concept}/{reference}` (European Commission)
- Specific Pattern: <http://www.irishstatutebook.ie/eli/2014/act/15/> (Irish Statute Book)

International research<sup>6</sup> indicates that the elements under consideration for inclusion in the National URI Pattern should include:

#### 1. {domain} element

The {domain} component contains the Internet domain and, optionally, a path within that domain.

#### 2. {type} element

{type} indicates which kind of URI is involved. This may be:

- 'id' - identifier of an object (individual/instance) in a register.
- 'doc' - documentation (metadata) on the object in the register.
- 'def' - definition of a term in an ontology

#### 3. {concept} element

{concept} gives the human reader an indication of the type of concept that is identified by the URI.

#### 4. {reference} element

{reference} is the identifying name or code of the individual object.

Other possible elements include: **{namespace}** (For new URI sets placed under common governance), **{sector}** (Same categories that are included in data.gov.ie, and **{language}**

<sup>6</sup> E.g. [http://www.pilod.nl/w/images/a/aa/D1-2013-09-19\\_Towards\\_a\\_NL\\_URI\\_Strategy.pdf](http://www.pilod.nl/w/images/a/aa/D1-2013-09-19_Towards_a_NL_URI_Strategy.pdf)

## Annex 1: Open Data Glossary

This is a general Glossary of Open Data Terms and Acronyms, for use as a reference guide for the Open Data Initiative. This Glossary will be expanded and enhanced as required.

### Open Data

Data broadly refers to information, rendered in a human- or computer-readable manner, which may be the subject of research or a raw product of research. A dataset may be considered Open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike. (Open Knowledge Foundation - [opendefinition.org](http://opendefinition.org))

### Open Government Data

Data which has been produced or gathered by public bodies during the course of business activities, and published under an Open Licence.

### Data Protection

Data protection legislation protects privacy rights of individuals in relation to the processing of their personal data.

### Anonymisation and Aggregation

Anonymisation and aggregation can be used to ensure that datasets relating to human subjects comply with relevant data protection legislation before publication.

Anonymisation involves the redaction of information from a dataset where individuals could previously have been identified.

Aggregation involves the publication of a dataset in summary form to exclude personal information which would allow an individual to be identified.

Data Protection legislation applies to the publication of datasets as Open Data. Where datasets intended for publication contain personal data they should be effectively anonymised in an irreversible manner that does not allow an individual to be re-identified, singled-out or inferred, in which case it will not be considered to be personal data.

### Copyright

Copyright is an area of Intellectual Property law which covers original creative works including literary, dramatic, musical and artistic works, film, sound recordings, broadcasts and the typographical arrangement of published editions, computer software and non-original databases, and performances. Copyright exists from the moment the work is created, and does not require any registration of the work. In the case of a dataset, copyright may belong to an employer, a government department, a funder or another party, depending on the contract surrounding the creation of the data.

### Licence

Licensing allows copyright owners to permit approved use and reuse of their work, without relinquishing copyright fully. Licensing can permit both commercial and non-commercial reuse of a work, depending on the terms of the licence, and licences may last in perpetuity or for a specified period. The application of a licence does not mean that a copyright statement should not be applied to a work, and many licences such as Creative Commons suggest that the copyright holder is credited. Open Data is usually associated with an Open Licence such as CC-BY (Creative Commons Attribution Only) or a Publication Domain Dedication such as CC0.

## CC0

CC0, Public Domain Dedication or “No Rights Reserved”

(<https://creativecommons.org/publicdomain/zero/1.0/>) is not truly a Creative Commons licence, as it does not reserve any rights in a copyright work. Assigning a Public Domain Dedication to a work relinquishes all rights in it, and allows use and re-use of the work for any purpose, without credit to the original author. Essentially this dedication allows works to enter the public domain before the legal term of copyright protection has ended. A Public Domain Dedication cannot be revoked.

## Attribution Licence

A licence requiring that the original source of the licensed material is cited (attributed).

## CC-BY (Creative Commons Attribution)

This licence lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This licence is recommended for maximum dissemination and reusability of licensed materials.

## Machine-readable format

Machine-readable formats are those containing structured data which can be extracted and analysed in an automated way. Examples of machine readable formats include those with a tabular structure such as .xls and .csv, as well as formats such as XML and JSON which are more flexible.

Generally, popular human-readable formats such as Word documents, PDF and HTML include formatting and display information which means that they are not machine-readable.

## Open formats

An open format is one where the specifications for the software are available to anyone, free of charge, so that anyone can use these specifications in their own software without any limitations on reuse imposed by intellectual property rights. Open formats include .csv and .xml.

## Proprietary formats

A format is proprietary if it encodes data in so that a file is readable only by using the same type of software used to create the file. Proprietary software does not openly publish its specifications for reuse. Proprietary formats include .xls (created in Microsoft Excel) and .docx (created in Microsoft Word).

## Application Programming Interface

An Application Program Interface (API) is a set of routines, protocols, and tools for building software applications. Programs that use a common API will have similar user interfaces, making it easier for users to learn new programs. (<http://www.w3.org/egov/wiki/Glossary>)

## Commonly used acronyms

ASCII	American Standard Code for Information Interchange
CC	Creative Commons
CKAN	Comprehensive Knowledge Archive Network
CSO	Central Statistics Office of Ireland
CSV	Comma Separated Value
DCAT	Data Catalogue
DCAT-AP	DCAT Application Profile
DCMI	Dublin Core Metadata Initiative
EC	European Commission
ETR	European Terrestrial Reference

GeoJSON	Geo JavaScript Object Notation
GML	Geography Markup Language
GTFS	<i>General Transit Feed Specification</i>
HIQA	<i>Health Information and Quality Authority</i>
IATI	International Aid Transparency Initiative
IETF	Internet Engineering Task Force
IFC	International Finance Corporation
INSPIRE	Infrastructure for Spatial Information in the European Community
ISO	International Organization for Standardization
ITM	Irish Transverse Mercator
JSON	JavaScript Object Notation
KML	<i>Keyhole Markup Language</i>
LAS	Log ASCII Standard
NetCDF	Network Common Data Form
OASIS	Organization for the Advancement of Structured Information Standards
ODF	OpenDocument Format
OGC	Open Geospatial Consortium
OSi	Ordnance Survey Ireland
PBWG	Public Bodies Working Group
PDF	<i>Portable Document Format</i>
PSI	Public Sector Information
RDF	Resource Description Framework
SDMX	Statistical Data and Metadata eXchange
URI	Uniform Reference Identifier
W3C	World Wide Web Consortium
WGS	World Geodetic System
WHO	World Health Organization
WKT	Well Known Text
WFS	Web Feature Service
WMS	Web Map Service
XBRL	eXtensible Business Reporting Language
XML	Extensible Markup Language

## Annex 2: Open Data Checklist

### Open Data Publication Planning

Dataset Details		
<b>Organisation Responsible</b>		
<b>Open Data Audit Date</b>	Select	
<b>Dataset Name</b>	Select	
<b>Dataset Description</b>		
<b>Does your organisation produce, manage and maintain the dataset?</b>	Select	
<b>Is this data published, online or offline?</b>	Select	Please provide link if already online:
<b>Is the dataset updated?</b>	Select	How frequently? Select
<b>If frequency of update is "other", please provide details.</b>		

<b>Is there a demand for this data through FOI, PSI, AIE?</b>	Select	
<b>Is this dataset associated with a licence?</b>	Select	Which licence?
<b>Publication</b>		
<b>Is anonymisation and/or aggregation required (if, for example, the dataset contains personal or commercially sensitive data?)</b>	Select	Is there a reason why this cannot be done? (Please explain why):
<b>Is there any other reason why data cannot be published (eg., Confidentiality clauses, third party copyright, etc)?</b>	Select <input type="text"/>	Please provide detail on reason:
<b>Risks &amp; Issues</b>		
<b>Publish Y/N</b>	Select	If, No specify Reason
<b>Open Data</b>		
<b>Can this dataset be associated with the recommended Open Licence (CC-BY)</b>	Select	
<b>Will the Open Licence be used within 6 months?</b>	Select	

<b>Is this dataset published in structured, machine-readable format?</b>	Select
<b>Is this dataset published in a non-proprietary format (in addition to PDF, Excel or Word)?</b>	Select
<b>Will the Non-Proprietary Format be used within 6 months?</b>	Select
<b>Is this dataset associated with the Open Data Technical Framework (DCAT-AP) metadata profile?</b>	Select
<b>Will the metadata profile be used within 6 months?</b>	Select
<b>Does your dataset use a recognised International or National Standard, as outlined in the Technical Framework?</b>	Select
<b>Which Standard?</b>	
<b>What are the potential benefits associated with publishing this dataset as Open Data?</b>	
<b>What is the Estimated Total Cost of delivering as Open Data?</b>	Total Cost: <input type="text"/> Calculation Basis:
<b>If this dataset is to be published as Open Data, what is the estimated publication timescale?</b>	

Open Data Candidate	Select	If, No specify Reason
<b>Status and Actions</b>		
<b>Updates</b>		
<b>Actions</b>		
<b>Attachments</b>		