

Office of the CIO

Open Health Data Policy



June 2016
Version 1.0

101100011

Title:	Open Health Data Policy
Purpose:	To define the required governance, preparation and publication of Open Data within the HSE.
Author:	Office of the Chief Information Officer, HSE.
Publication date:	June 2016
Target Audience:	All HSE data users, including HSE staff, contractors, and the general public.
Superseded Documents:	-
Related Documents:	HSE Information Technology (I.T.) Security Policy HSE Information Technology Acceptable Use Policy HSE Information Classification & Handling Policy HSE Access Control Policy
Review Date:	
Contact Details:	

1	Open Health Data Vision.....	5
2	Purpose.....	5
3	Scope	5
4	Open Health Data Principles	5
5	Definitions.....	6
6	Glossary.....	7
7	Related Legislation and Policy	7
	International	7
	National.....	8
	HSE-Specific.....	8
8	Policy.....	9
	Open Data Programme	9
	Open Data Internal Communication	9
	Open Data Public Engagement	9
	Open Data Inventory.....	9
	Data Security & Protection.....	10
	Open Data Formats, Identifiers and Vocabularies	10
	Open Data Classification.....	10
	Open Data Metadata	10
	Open Data Licencing	10
	Open Data Quality.....	10
	Open Data Quality Assessment.....	11
	Open Data Cleansing	11
	Open Data Access and Sharing.....	11
	Open Data Usage.....	11
	Open Data Archiving and Retention.....	11
9	Roles & Responsibilities.....	11
	Office of the Chief Information Officer (OCIO), HSE	11
	Open Health Data Programme Manager.....	11
	Data Stewards	12
	Open Health Data Oversight Authority.....	12



Staff, Contractors, Sub-contractors, Agency Staff, and Third Party Commercial Service
Providers..... 12

1 Open Health Data Vision

Improve patient care, empower choice and support innovation through open access to HSE data, while protecting patients' personal data.

2 Purpose

The purpose of this Open Health Data Policy is to create an overarching framework for the eHealth Ireland Open Data for Health Strategic Programme¹. This policy defines the required governance, preparation and publication of Open Data within the HSE.

The Open Health Data Policy will help to ensure consistency of approach to the release of Open Data across the HSE, aligned with national guidelines and international best practice. Unless otherwise agreed with a full rationale, it is mandatory that all HSE data producers and consumers abide by this policy, and that all Open Data development and activities conform to this policy and the agreed Data Principles.

3 Scope

This policy represents the HSE's national position and takes precedence over all other relevant policies which are developed at a local level. The policy applies to:

- All HSE data producers, managers and consumers, including HSE staff, contractors, sub-contractors, agency staff and authorised commercial service providers.

4 Open Health Data Principles

The definition of core Open Health Data Principles ensures that there is a common set of values for Open Data across the HSE. The Principles are structured incrementally, building on each other in terms of prioritisation and remit. This lays the foundations for increasingly more valuable Open Data for the HSE itself, citizens, businesses, researchers and other Public-Sector Bodies. The Open Health Data Principles support the HSE's vision to 'improve patient care, empower choice and support innovation through open access to HSE data, while protecting patients' personal data'.

- Data is a valuable asset to improve healthcare delivery and planning
- Personal data is protected under data protection legislation
- Open Health Data is proactively published
- Open Health Data is usable by all
- Open Health Data is of high quality
- Reuse of Open Health Data will be supported
- Researchers and other external stakeholders have access to relevant data under controlled circumstances.
- People can access data about themselves and how it is being used.

¹ <http://www.ehealthireland.ie/Strategic-Programmes/Open-Data-for-Health/>

5 Definitions

A list of terms used throughout this policy.

Confidential Information: Confidential information is defined as information which is protected by Irish and/or E.U. legislation or regulations, HSE policies or legal contracts. The unauthorised or accidental disclosure of this information could adversely impact the HSE, its patients, its staff and its business partners².

Data Steward: The individual responsible for the management of a business area within the HSE.

Dataset: A named collection of related records, with the collection containing data organised or formatted in a specific or prescribed way, often in tabular form.

Internal Information: Internal information is defined as information that is only intended for internal distribution among HSE staff, students, contractors, sub-contractors, agency staff and authorised third parties (i.e. service providers etc.). In the majority of instances there would be no significant impact on the HSE, its staff, clients or patients if this type of information was mishandled or accidentally released².

Open Data: Data that anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)³.

Open format: Any widely-accepted, non-proprietary, platform-independent, machine-readable method for transmitting data, which permits automated processing of such data and facilitates analysis and search capabilities.

Public Information: Public information is defined as information that is available to the general public and is intended for distribution outside the HSE. There would be no impact on the HSE, its staff, clients or patients if this type of information was mishandled or accidentally released².

Restricted Information: Restricted information is defined as highly sensitive confidential information. The unauthorised or accidental disclosure of this information would seriously and adversely impact the HSE, its patients, its staff and its business partners².

Third Party Commercial Service Provider: Any individual or commercial company that have been contracted by the HSE to provide goods and/or services (for example, project / contract management, consultancy, information system development and/or support, supply and/or support of computer software / hardware, equipment maintenance, information management services, patient / client care and management services etc.) to the HSE.

² http://www.hse.ie/eng/services/publications/pp/ict/Information_Classification_Handling_Policy.pdf

³ <http://opendefinition.org/>

6 Glossary

API	Application Programming Interface
CC	Creative Commons
CKAN	Comprehensive Knowledge Archive Network
CSO	Central Statistics Office of Ireland
CSV	Comma-Separated Values
DCAT	Data Catalog Vocabulary
DCAT-AP	DCAT Application Profile
DPER	Department of Public Expenditure and Reform
JSON	JavaScript Object Notation
NHS	UK National Health Service
OECD	Organisation for Economic Co-operation and Development
PSI	Public Sector Information
RDF	Resource Description Framework
URI	Unique Reference Identifier
URL	Unique Reference Locator
W3C	World Wide Web Consortium
XML	Extensible Markup Language

7 Related Legislation and Policy

International

There is international legislation and a number of best practices in relation to the publication and use of Open Data and Open Health Data. These include but are not limited to:

- Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information⁴
- W3C's Data on the Web Best Practices⁵
- The Sunlight Foundation's Open Data Policy Guidelines⁶
- OECD Health Data Governance review on Open and Transparent Health Information Systems⁷
- The GovLab and NHS England's study on The Open Data Era in Health and Social Care⁸

⁴ <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013L0037>

⁵ <https://www.w3.org/TR/dwbp/>

⁶ <http://sunlightfoundation.com/opendataguidelines/>

⁷ <http://www.oecd.org/publications/health-data-governance-9789264244566-en.htm>

⁸ <http://thegovlab.org/nhs/>

National

The HSE has an obligation to abide by all relevant Irish legislation and European legislation. The relevant acts, strategies and policies, which apply to the Open Health Data Policy, include but are not limited to:

- National Open Data Strategy⁹
- National Open Data Technical Framework¹⁰
- Circular 12/2016: Licence for Re-Use of Public Sector Information¹¹
- European Communities (Re-use of Public Sector Information) (Amendment) Regulations 2015¹²
- Freedom of Information Act 2014¹³
- The Data Protection Acts (1988/2003)¹⁴
- European Communities (Electronic Communications Networks and Services) (Data Protection and Privacy) (Amendment) Regulations 2008¹⁵
- Data Sharing and Governance Bill¹⁶
- The Health Research Board's Proposals for an Enabling Data Environment for Health and Related Research in Ireland¹⁷

HSE-Specific

A list of the HSE National IT Policies and Standards are available on the HSE website¹⁸. The relevant policies and standards, which apply to the Open Health Data Policy, are:

- HSE Information Security Policy
- HSE Information Technology Acceptable Use Policy
- HSE Access Control Policy
- HSE Information Classification & Handling Policy
- HSE Standard Terms For ICT Services and Supplies
- HSE Knowledge and Information Plan¹⁹
- HSE eHealth Strategy for Ireland²⁰

⁹ <https://data.gov.ie/strategy>

¹⁰ <https://data.gov.ie/technical-framework>

¹¹ <http://circulars.gov.ie/pdf/circular/per/2016/12.pdf>

¹² <http://www.irishstatutebook.ie/eli/2015/si/525/>

¹³ <http://www.irishstatutebook.ie/eli/2014/act/30/enacted/en/>

¹⁴ <http://www.irishstatutebook.ie/eli/2003/act/6/enacted/en/>

¹⁵ <http://www.irishstatutebook.ie/eli/2008/si/526/made/en>

¹⁶ <http://www.per.gov.ie/en/datasharing/>

¹⁷ <http://www.hrb.ie/publications/hrb-publication/publications//709/>

¹⁸ <http://www.hse.ie/eng/services/publications/pp/ict/>

¹⁹ <http://ehealthireland.ie/Knowledge-Information-Plan/>

²⁰ <http://ehealthireland.ie/Knowledge-Information-Plan/eHealth-Strategy-for-Ireland.pdf>

8 Policy

Open Data Programme

1. Promote the proactive release of 'Public' data as Open Health Data, so that anyone can freely access, use, modify, and share it for any purpose (subject, at most, to requirements that preserve provenance and openness).
 - a. 'Public' data refers to any data that is, or can be, classified as 'Public' according to the HSE Information Classification scheme, as defined in the HSE Information Classification & Handling Policy².
2. Minimise limitations on the disclosure of public information while appropriately safeguarding internal, confidential and restricted information, according to HSE Information Classification.
3. Where possible, provide access to free, historical archives of all Open Health Data.
4. Maintain Open Health Data on the National Open Data Portal, <https://data.gov.ie>, so that people can discover and access the datasets.
5. When publishing any report that includes tables, publish the accompanying raw data as Open Health Data along with the document.
6. Include requirements for Open Health Data in all procurement practices.
7. Assign appropriate funding to achieve the goals of this programme.

Open Data Internal Communication

8. Provide internal awareness campaign on the importance and benefits of publishing Open Health Data
9. Proactively consult with HSE staff on the implementation and ongoing development of this policy.
10. Provide training for Data Stewards and all HSE staff on their roles and responsibilities in the Open Health Data Programme.

Open Data Public Engagement

11. Maintain a dedicated online location, for example on the eHealth Ireland website, where people can stay informed about the Open Health Data Programme, including access to this Policy, minutes, guidelines, presentations, progress updates, evaluation reports, event notifications, etc.
12. Proactively consult with the public on the implementation and ongoing development of this policy.
13. Engage with the public on what datasets will have the greatest benefit if published as Open Health Data.
14. Support innovative uses of the Open Health Data.
15. Gather and make available feedback from data users on published datasets and on the Open Health Programme as a whole.

Open Data Inventory

16. Start compiling a public inventory of datasets held within each HSE organisation.
17. Provide an explanation for datasets that are not made available as Open Health Data.

18. Specify methods of determining the prioritisation of data release, which takes into account new and existing signals of interest from the public, the HSE's programmatic priorities, existing opportunities for data use in the public interest, and cost.

Data Security & Protection

19. Open Health Data will not infringe upon a person's right under data protection legislation and policies, outlined in section 7.
20. Safeguarding internal, confidential and restricted information will be maintained as a paramount priority.
21. Develop and implement a process for determining the relative public benefit and cost associated with statistically de-anonymising and/or aggregating internal, confidential and restricted information, so as it can be safely published as Open Health Data.
22. Align with the Health Research Board's recommendations on publishing Health Research Data in a safe manner. Health Research Data may or may not be published as Open Health Data, provided that it does not infringe on data protection legislation.

Open Data Formats, Identifiers and Vocabularies

23. Publish Open Health Data in machine-readable, non-proprietary, standardised data formats, including but not limited to CSV, JSON, XML and RDF.
24. If available and practical, provide data in multiple formats
25. Use persistent Unique Resource Identifiers (URIs) as identifiers of datasets
26. Use persistent URIs as identifiers within datasets
27. Use URIs or terms from existing vocabularies and codelists where possible.

Open Data Classification

28. Data should be classified as 'Public' according to the HSE Data Classification scheme² before being published as Open Health Data.

Open Data Metadata

29. Provide descriptive and structural metadata in a standardised machine-readable format, namely the W3C Data Catalog Vocabulary (DCAT)²¹ and more specifically, the DCAT Application Profile for European Data Portals (DCAT-AP)²².

Open Data Licencing

30. Publish Open Health Data under an Open Licence, namely the Creative Commons Attribution 4.0 International (CC BY 4.0)²³ as recommended by Circular 12/2016: Licence for Re-Use of Public Sector Information¹¹. This signifies there are no restrictions on using, sharing and adapting the data for non-commercial or commercial purposes, other than attribution to the original data source.
31. Recommend the following attribution statement for data users:
 - a. *"Contains HSE Public Sector Information licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence".*

Open Data Quality

32. Publish high quality Open Health Data, i.e. data that is accurate, available, complete, conformant, consistent, credible, processable, relevant and up-to-date.
33. Provide data provenance information, i.e. the origin and history of the dataset.

²¹ <http://www.w3.org/TR/vocab-dcat/>

²² https://joinup.ec.europa.eu/asset/dcat_application_profile/home/

²³ <https://creativecommons.org/licenses/by/4.0/>

34. Provide an indication of the version number or date with the data, along with a complete version history.

Open Data Quality Assessment

35. Establish processes for ensuring that datasets are reviewed for use-appropriate format, high quality, timeliness, and exclusion of protected and sensitive information

Open Data Cleansing

36. Improve the value and processability of data by enriching raw data. Missing values can be filled in, the data structured according to standard formats and modelled according to standard schemas, and new attributes and measures can be added.

Open Data Access and Sharing

37. Provide bulk download of datasets so that users can access the complete dataset in one go.
38. Provide subsets for large datasets making it more convenient for users to store and process the data.
39. Make data available through an API, offering the greatest flexibility and processability for users of the data. APIs can enable real-time data usage, filtering on request, and the ability to work with the data at an atomic level.
 - a. Use REST and Web Standards as the foundation of APIs
 - b. Provide complete documentation for the API
 - c. Avoid changes to the API that break client code, and communicate any changes in the API to your developers when evolution happens.

Open Data Usage

40. Continuously monitor and report on the access and usage of Open Health Data, through technical analytical means, in addition to qualitative user engagement.

Open Data Archiving and Retention

41. Archive Open Data according to the organisation's data archiving policy, facilitating permanent, lasting access to data.

9 Roles & Responsibilities

Office of the Chief Information Officer (OCIO), HSE

The OCIO is responsible for:

- Approving and publishing the policy.
- The annual review of policy.
- Approving all changes and amendments to the policy.

Open Health Data Programme Manager

The Open Health Data Programme Manager is responsible for:

- The implementation of this policy within the HSE.
- Reporting to the OCIO on the progress and the implementation of the policy.
- Liaising with and advising data stewards on the implementation of the policy within their business area.
- Liaising with staff, contractors, sub-contractors, agency staff, and third party commercial service providers on their responsibilities under the policy.
- Liaising with the public on the publication of high-value Open Health Data and supporting its innovative use.

- Managing the publication of Open Health Data on the National Open Data Portal, <https://data.gov.ie>.
- Creating a roadmap to support the ongoing implementation of the Open Health Data Policy across the HSE
- Making sure adequate procedures and training programs are implemented to ensure on-going compliance of this policy.
- Preparing and publishing an annual report, including an assessment of progress towards achievement of the goals of the Open Health Data Policy, and a description and publication timeline for datasets envisioned to be published by the HSE in the following year.

Data Stewards

Data Stewards are responsible for:

- The implementation of this policy within the business areas for which they are responsible.
- Reporting to the Open Health Data Programme Manager on the progress and the implementation of the policy.
- Ensuring all data (irrespective of its format) owned, created, received, stored and processed within the HSE business area they manage that is classified as 'Public' is handled in accordance with this policy.
- Ensuring that all HSE staff, students, contractors, sub-contractors and agency staff who report to them are made aware of, understand and have access to this policy, are instructed to comply with this policy, and are provided with adequate information and training regarding the implementation of this policy.

Open Health Data Oversight Authority

The Open Health Data Oversight Authority is responsible for:

- Measuring the adoption of the Open Health Data Policy across the HSE.
- Measuring the impact of the publication of datasets from the HSE, by connecting Open Health Data publication and usage to the Business Cases outlined in the previous section.
- Mandating a review for potential changes to this policy.

Staff, Contractors, Sub-contractors, Agency Staff, and Third Party Commercial Service Providers

Each staff member, contractor, sub-contractor, agency staff member, and third party commercial service provider is responsible for:

- Complying with the terms of this policy.
- Reporting all misuse and breaches of this policy to their line manager.

