Open Health Data Governance Strategy
The publication of Irish Public Sector data in open, free and reusable formats is recognised as a key element of the Public Service Reform agenda. Through the national (http://data.gov.ie) and regional Open Data Portals, citizens can access raw Public Sector data, providing insights into the operation of public organisations and services. The aim of Open Data is twofold; on the one hand facilitating transparency of the Public Sector and on the other providing a valuable resource that can drive innovation.

The availability of Open Data can empower citizens and support clinicians, care providers, and researchers make better decisions, spur new innovations and identify efficiencies while ensuring that personal data remains confidential.

Richard Corbridge, CIO HSE
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 Governance Strategy is to optimise the value of Open Data published by the HSE so that the Open Health Data Vision is achieved. The Open Health Data Governance Strategy ensures the consistent and reliable preparation, publication and maintenance of Open Health Data across all arms of the HSE, through a standard set of policies, best practices, and structured processes.

3 Scope

This strategy applies to:

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

4 Introduction to Open Health Data

In October 2013, Minister for Public Expenditure and Reform, Brendan Howlin T.D., announced the Open Data Initiative for Ireland to promote transparency and inspire innovation. The National Open Data Portal data.gov.ie was launched in May 2014.

The adoption of Open Data has many benefits, including improved provision of services, increased transparency and access to public information, and enhanced coordination and efficiencies within and among departments, partner organisations and citizens.

A recent McKinsey report on ‘Open data: Unlocking innovation and performance with liquid information’ estimates the use of open and proprietary data in health care could help generate value of $300 billion to $450 billion per year in the United States alone. They assert most of this value comes in the form of cost savings to providers, payers, and patients. A study by the GovLab at New York University and NHS England on ‘The Open Data Era in Health and Social Care’ states that Open Data holds particular potential in the health sector. By releasing health data to patients and, when appropriate, on an anonymized basis to researchers and the public, governments and healthcare organizations are betting on the power of greater openness of data to improve the quality of care, lower healthcare costs, and facilitate patient choice.

An OECD report on ‘OECD Health Data and Data Governance’ in 2015 highlighted that on the one hand, health information systems are developed for the public’s benefit, to ensure health care is accessible, of high quality, and affordable and to generate new scientific discoveries to improve therapies, outcomes and overall population health. However, on the other hand, public awareness of

---

1 http://sunlightfoundation.com/opendataguidelines/
the data inputs to the health information system, the users of the system, the uses of the system and
the benefits and risks that are associated with the system is often limited (OECD HCQI Expert Group
2015).

The OECD report also looked at whether countries participating in the study had a policy or a
programme in place to promote open government health data. It found that while Ireland reported an
e-government strategy for 2012 to 2015 to make better and more innovative use of information and
communications technologies to improve the public-sector customer experience including data
accessibility; the initiative does not include health data.

Open Data for Health has been prioritised as a Strategic Programme within the Office of the CIO.4

5 Current Data Publication Practices in the HSE

The HSE is data rich, with large amounts of data being collected. HIQA’s Catalogue of National
Health Information Sources lists 107 separate data collections. Many of these information sources
contain personal data that are not suitable for publication. Some data is released under a controlled
environment and licenced for particular purposes, for example, data used by researchers governed by
organisations such as the Health Research Board. It is critical that these data protection and data
governance measures are in place, so that the protection of privacy, confidentiality and security is
maintained as a paramount priority.

In this context the HSE’s Open Health Data Initiative can also advance. The following data is suitable
for publication as Open Data: (i) data that does not pertain to individuals so does not contain personal
information, or (ii) data that has been de-identified or aggregated using standard statistical methods
so that it no longer contains personal information that would infringe on data protection rights.

There are currently examples of health data being published in the public domain, including monthly
HSE performance reports, the Health Atlas, and HIPE annual reports. Additionally, most HSE units
publish regular reports, publications and statistics on their websites that contains valuable data.

The current strengths of the HSE in relation to the Open Data for Health Strategic Programme are:

- The HSE is data rich. A huge amount of data is produced and used daily.
- Much of this data is structured and coded using national and international best practices. Approaches to coding are carefully documented.
- Reporting mechanisms are already in place that export data from a variety of systems and
  present the data in human-readable formats. These processes can be utilised to export the data in
  machine-readable formats as well.
- Data aggregation methods are in use to convert identifiable, personal data to statistical,
  non-identifiable data. Aggregate, statistical data is suitable for publication as Open Data
  without infringing on a person’s rights under data protection legislation.

The challenges of the Open Data for Health Strategic Programme are:

- There is an abundance of data in the HSE that could be safely published as Open Health
  Data to achieve the benefits outlined below, but is not currently being made available.
- Data is published in pdf and word reports, or in web applications, such as Health Atlas.
  This means the data is not easy to access for processing and analysis purposes.
- There is no discovery mechanism for health data; a user must know what data they are
  looking for, who publishes the data and where the data is published.

---

4 [http://www.ehealthireland.ie/Strategic-Programmes/Open-Data-for-Health/](http://www.ehealthireland.ie/Strategic-Programmes/Open-Data-for-Health/)
5 [https://www.hiqa.ie/healthcare/health-information/data-collections/online-catalogue](https://www.hiqa.ie/healthcare/health-information/data-collections/online-catalogue)
6 [http://www.hrb.ie/](http://www.hrb.ie/)
8 [http://www.hse.ie/eng/services/maps/](http://www.hse.ie/eng/services/maps/)
9 [http://www.hpo.ie/HIPE_and_NPRS_Reports.htm](http://www.hpo.ie/HIPE_and_NPRS_Reports.htm)
- Health data is often shared on an ad-hoc basis, in response to a request from another Public Body or a member of the public. This is time-consuming and can lead to inconsistent reporting.
- There is currently not an established role for Open Data management within the Health system. There are often poor data maintenance and standardisation techniques in place.

6 The Data Spectrum

![The Data Spectrum](https://theodi.org/data-spectrum)

The Open Data Institute in the UK created The Data Spectrum to clarify the differences between Closed, Shared and Open Data. **Closed Data** is data that is only used internally in an organisation or team. **Shared Data** is data that is shared with certain people or groups under certain, controlled conditions. **Open Data** is data that anyone can freely access, use, modify, and share for any purpose, (subject, at most, to requirements that preserve provenance and openness.)\(^\text{11}\). Whether big, medium or small, whether state, commercial or personal, the important thing about data is how it is licensed.

Examples of HSE data that is categorised according to The Data Spectrum is as follows:

- **Internal access**: Patients personal health record
- **Named access**: Prescription data shared with a named pharmacy
- **Group-based access**: Patient treatment and outcome data for medical research
- **Public access**: Postcode service data
- **Anyone**: Performance reports, waiting times, available services

---

\(^{10}\) [https://theodi.org/data-spectrum](https://theodi.org/data-spectrum)

\(^{11}\) [http://opendefinition.org/](http://opendefinition.org/)
7 **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
- Reuses 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.

8 **Benefits of Open Health Data**

There are many examples internationally that demonstrate the positive impact of publishing Open Health Data. In the U.S., the Health Datapalooza national conference[^1] focuses on liberating health data, and bringing together the companies, start-ups, academics, government agencies, and individuals with the newest and most innovative and effective uses of health data to improve patient outcomes. The breadth of speakers, exhibitors and attendees reveals the potential scope of Open Health Data Initiatives.

Through a review of existing evidence and literature, the GovLab at New York University and NHS England have identified six domains that can be positively affected through open data initiatives in the health sector[^3]. These value propositions, depicted in Figure 2, are as follows:

- **Accountability**: The use of data to hold healthcare organizations and providers accountable for treatment outcomes.
- **Choice**: Providing open data to help patients make informed choices from among the healthcare options available to them.
- **Efficiency**: Improving the efficiency and cost-effectiveness of healthcare delivery.
- **Outcomes**: Improving treatment outcomes by using open data to make the results of different treatments, healthcare organizations, and providers' work more transparent.
- **Patient satisfaction and customer service**: Using open data to educate patients and their families and make healthcare institutions more responsive.
- **Economic growth and innovation**: Using open data to fuel new healthcare companies and initiatives, and to spur innovation in the broader economy.


Business Cases aligned with eHealth Ireland’s Knowledge & Information Plan

In this section, we look at a number of business cases that clearly articulate the business opportunities for the HSE in publishing Open Health Data.

eHealth Ireland’s Knowledge & Information Plan\(^{13}\) outlines five focus areas which are key to facilitating the seamless delivery of healthcare across integrated care pathways and all stages of care, identified in consultation with the HSE National Directors and clinicians. These are:

- Care Delivery Enablement
- Electronic Health Records
- Cross Setting Information Integration
- Health Service Insights
- National Support Systems

For each of these focus areas, we present a business case that will demonstrate how the publication of Open Health Data can positively impact this capability area.

**Care Delivery Enablement**

This capability delivers a series of clinical and care delivery capabilities which digitise areas/processes thereby allowing electronic data capture, better quality of care, more clinician time spent with patients and greater ability for patients to participate in their own care.

Care Delivery Enablement, by its very nature, results in the production of vast quantities of data. While a lot of this data is personal information related to particular patient cases, de-identified and aggregated care delivery data can be published as Open Health Data. For example the volume of patients being treated in particular programmes, the time spent under care, the number of eReferrals and average outcomes can be extrapolated.

An example of care delivery data that is currently published is the HIPE ‘Activity in Acute Public Hospitals in Ireland’ reports\(^{14}\). This data includes information about hospital discharges per patient type, hospital type, diagnoses, etc.

\(^{13}\) [http://www.ehealthireland.ie/Knowledge-Information-Plan/](http://www.ehealthireland.ie/Knowledge-Information-Plan/)

\(^{14}\) [http://hpo.ie/](http://hpo.ie/)
Benefits of Open Health Data from the Care Delivery Enablement focus area

- Improved decision-making within the HSE and across the public-sector based on real data.
- More efficient care delivery services, optimised through the analysis of current practices.
- Transparency around current care delivery programmes.

Electronic Health Records

This capability creates and enriches the patients’ electronic health records at each setting, stores the detailed care information at activity level, and feeds the summary care record to enable cross setting integration.

Similarly to care delivery, Electronic Health Records contain personal information pertaining to individuals. However, de-identified and aggregated data around electronic health records and their surrounding processes can be published as Open Health Data. For example, the numbers of generic vs branded medication being prescribed nationally, or types of care plans being utilised.

Benefits of Open Health Data from the Electronic Health Records focus area

- Contributes to a better understanding of service utilisation.
- Safer and more effective patient care by looking at overall trends.

Cross Setting Information Integration

This capability delivers the required integration, information flows and process standardisation across care settings that reflect the patient care pathway, enabling the summary care record at a patient level, seamless transition between settings supported by secure handovers of patient information.

Publishing Open Health Data related to Cross Setting Information Integration provides an insight into how increased collaboration and data interoperability improves patient care. For example, what is a complete listing of available services, what standards are being used for data exchange, and how are information flows happening in practice, for example, eReferrals from GPs to hospitals.

Benefits of Open Health Data from the Cross Setting Information Integration focus area

- Evidence of robust security around information sharing practices.
- Increased transparency around information flows within the HSE.
- Greater efficiencies in data access by data practitioners.

Health Service Insights

This capability delivers the information management, reporting and analysis solutions and processes which provide timely, reliable information and decision support for patients, clinicians, and management from the micro level (e.g. individual patients or treatments) up to the macro view (e.g. system wide performance, population health trends).

Much of the macro-view Health Service Insights are suitable for publication as Open Health Data, as they do not contain personal or sensitive data. These insights are of direct benefit to health care professionals, but also to a wider audience across the public sector, business and general public. These health service insights offer an overview of processes, systems, patient care, etc. and can be used to identify opportunities for improvement, as input for future decision-making, or to support new innovations. For example, the HSE currently publishes monthly performance reports.

Benefits of Open Health Data from the Health Service Insights focus area

- Increased understanding of the sector as a whole through advanced business intelligence performed by multiple stakeholders, not only internal data scientists.
- New innovations based on insights gleaned through a wide-range of data analytics.
- Enabling the HSE to adopt innovative approaches.
National Support Systems

This capability provides the national backbone for the core support service – finance, procurement, and HR – it is crucial for integrated, efficient management of the health system as a whole.

Open Data from the HSE support systems includes administrative information on overall staffing, procurement and expenditure, which gives valuable insights to the HSE and beyond on the operations of the organisation.

Benefits of Open Health Data from the National Support Systems focus area

- Transparency of the overall running of the HSE, including financial, procurement and HR.
- Increased accountability around internal decision-making.
- Ability to demonstrate impact of cost-savings measures.

9 Flagship Projects

The purpose of the flagship projects is to demonstrate the actual business value that can be achieved through the publication of Open Health Data, which will ultimately lead to the realisation of the HSE Open Health Data Vision to:

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

The HPO and HealthAtlas datasets can support improved patient care, as more knowledge is available to patients on what services are available, in addition to details of particular services in the form of in-patient and day patient discharges from acute public hospitals. Open Health Data from Healthlink informs people on the scale of messaging and referrals that takes place between Hospitals, Health Care Agencies and General Practitioners. Having information on health services from HealthAtlas available as Open Health Data, empowers people to be able to make healthcare choices that are right for them. To encourage new innovations in healthcare is important to advance patient care. The publication of information such as the Primary Care Reimbursement Service (PCRS) data as Open Health Data provides innovators with valuable insights into current HSE practices.

Healthcare Pricing Office (HPO)

From 1st January 2014 the National Casemix Programme and the Health Research and Information Division at the ESRI became the Healthcare Pricing Office (HPO)\(^{15}\). The HPO is responsible for the management of the Hospital In-Patient Enquiry Scheme (HIPE) and the National Perinatal Reporting System (NPRS).

HIPE is a system that collects information on hospital day cases and in-patients in Ireland. They provide training in clinical coding to all hospital staff working in this area, in addition to developing and supporting the relevant software requirements. NPRS provides national statistics on perinatal events, in particular data on pregnancy outcomes, perinatal mortality and important aspects of perinatal care.

The HPO Open Health Data Flagship Project involves the publication of 74 HPO datasets, based on the ‘Activity in Acute Public Hospitals in Ireland Annual Report 2014’\(^ {16} \). The Activity in Acute Public Hospitals in Ireland Annual Report, 2014, is a report on in-patient and day patient discharges from acute public hospitals participating in the Hospital In-Patient Enquiry (HIPE) scheme in 2014. Discharge activity is examined by type of patient and hospital, and by demographic parameters (such as age and sex). Particular issues of relevance to the Irish health care system covered in the report relate to the composition of discharges by medical card and public/private status. Discharges are also analysed by diagnoses, procedures, major diagnostic categories, and diagnosis related groups. Maternity discharges are examined separately from other discharges. The analysis is presented at the national level. In 2014 HIPE discharges were coded using ICD-10-AM/ACHI/ACS 6th Edition and grouped into AR-DRG Version 6.0.

\(^{15}\) [http://www.hpo.ie/](http://www.hpo.ie/)

Healthlink

The National Healthlink Project\(^\text{17}\) provides a web-based messaging service which allows the secure transmission of clinical patient information between Hospitals, Health Care Agencies and General Practitioners. The project has been in operation since 1995 and has developed considerably since that time to its current status as national messaging broker. Healthlink are preparing datasets, such as the total number of outbound messages transmitted through Healthlink broken down by message type, to be published in open, machine-readable formats, based on their current reporting processes, so that they can be added to data.gov.ie. Statistical datasets include:

- GP Messaging
- Lab Orders
- Cancer Referrals
- Neurology Referrals
- General Referrals

HealthAtlas

Health Atlas provides details of available health services via a map interface\(^\text{18}\), namely:

- Health Centres
- General Acute Hospitals
- GP/Family Doctor
- Community Hospitals
- Pharmacy
- Dental Practice
- Nursing Homes
- GP Out of Hours

The Health Centres\(^\text{19}\) and General Acute Hospitals\(^\text{20}\) datasets were made available on data.gov.ie as Open Data under an Open Licence. The Health Intelligence Unit is preparing the other datasets to be published in open, machine-readable formats, so that they can be added to data.gov.ie.

Primary Care Reimbursement Service (PCRS)

The PCRS\(^\text{21}\) publications were reviewed for potential Open Data contained within the reports. The following datasets areas were identified:

- Schemes Overview
- Cardholder Data
- General Practitioner
- Pharmacy Data
- Dispensing Data
- Dental Data
- Optical Data

\(^{17}\) [http://www.healthlink.ie/](http://www.healthlink.ie/)
\(^{18}\) [http://www.hse.ie/eng/services/maps/](http://www.hse.ie/eng/services/maps/)
\(^{19}\) [https://data.gov.ie/dataset/health-centres-in-ireland](https://data.gov.ie/dataset/health-centres-in-ireland)
\(^{21}\) [http://www.hse.ie/eng/staff/PCRS/](http://www.hse.ie/eng/staff/PCRS/)
10 Roadmap

In order to align expectations of the various stakeholders involved in all aspects of data governance and to ensure those with a designated role fully understand their responsibilities, a RACI Matrix is used. RACI is a responsibility assignment matrix that explicitly outlines who is Responsible, Accountable, Consulted and Informed for each task. For each task, there should only be one assigned Responsible role and one Accountable role. There can be multiple Consulted and Informed roles.

- The Responsible role is the main doer of the task.
- The Accountable role is ultimately answerable for the delivery of the task.
- The Consulted role participates in communication on the task.
- The Informed role is kept up-to-date on progress of the task.

<table>
<thead>
<tr>
<th>Open Data Programme</th>
<th>Time</th>
<th>OCIO</th>
<th>OHD Programme Manager</th>
<th>Data Stewards</th>
<th>Flagship Project Data Steward</th>
<th>OHD Oversight Authority</th>
<th>Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appoint OHD Programme Manager</td>
<td>Q2 2016</td>
<td>R/A</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Appoint OHD Oversight Authority</td>
<td>Q3 2016</td>
<td>R/A</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Assign appropriate funding to achieve the goals of this programme.</td>
<td>Q3 2016</td>
<td>R/A</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Provide internal awareness campaign</td>
<td>Q3 2016</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Provide training for Data Stewards and all HSE staff</td>
<td>Q3/Q4 2016</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Maintain OMD page on Health Ireland website</td>
<td>Ongoing</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Start compiling a public inventory of datasets held within HSE</td>
<td>Q2/Q3/Q4 2016</td>
<td>A</td>
<td>C</td>
<td>R</td>
<td>I</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Monitor and report on the access and usage of OHD</td>
<td>Ongoing</td>
<td>A</td>
<td>C</td>
<td></td>
<td>I</td>
<td></td>
<td>R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flagship Projects</th>
<th>Time</th>
<th>OCIO</th>
<th>OHD Programme Manager</th>
<th>Data Stewards</th>
<th>Flagship Project Data Steward</th>
<th>OHD Oversight Authority</th>
<th>Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carry out data audit</td>
<td>Q2/Q3 2016</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Identify datasets that could potentially be published as OHD</td>
<td>Q2/Q3 2016</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Ensure datasets to be published as OHD comply with the OHD Policy</td>
<td>Q2/Q3 2016</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Enrich any datasets that do not comply with the OHD Policy</td>
<td>Q2/Q3 2016</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Receive approval on the publication of the datasets as OHD on data.gov.ie</td>
<td>Q2/Q4 2016</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td>R</td>
<td>I</td>
</tr>
<tr>
<td>Publish the datasets as OHD on data.gov.ie, either as a one-off publication or as a periodic harvest</td>
<td>Q2/Q3 2016</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Monitor the usage of the OHD</td>
<td>Ongoing</td>
<td>A</td>
<td>C</td>
<td></td>
<td>I</td>
<td></td>
<td>R</td>
</tr>
<tr>
<td>Update and maintain the OHD as required</td>
<td>Ongoing</td>
<td>A</td>
<td>C</td>
<td></td>
<td>I</td>
<td></td>
<td>R</td>
</tr>
</tbody>
</table>