**Dataset Specification
Management Process Report**

Standardising Data for the Future

Document Control

Revision History

Version	Author	Date	Changes
V1 Draft	Theresa Barry/ Yvonne Coughlan	01/12/2021	Initial Draft
V2 Draft	Theresa Barry, Yvonne Coughlan, Cliona O'Donovan, Ann O'Donohoe	14/01/2022	Updated with Feedback from DSMP Subject Matter Experts
V3 Draft	Theresa Barry/ Yvonne Coughlan	25/01/2022	Enterprise Architecture Approval
V4	Theresa Barry/ Yvonne Coughlan	**28/01/2022**	

Review & Approval

Name	Versions Signed off	Sign off Date
DSMP	**V3**	**14/01/2022**
Enterprise Architecture	**V3**	**25/01/2022**

Background

An immense amount of data is collected on a regular basis about health and social care services in Ireland. Health is information-intensive, creating huge volumes of data every day. This data is used for many important purposes such as informing decision making, monitoring of diseases, planning of services, policy making, improving population health and for international reporting purposes. It is estimated that up to 30% of the total health budget may be spent one way or another on handling information, collecting it, looking for it, and storing it (1).

As per the Slaintecare Implementation Strategy (2019) ICT has the potential to be the biggest and most effective driver of change and improvement for better patient outcomes across the health system. In the future, a coherent suite of eHealth solutions will underpin and support our overall vision for integrated, patient-centred care, population health planning and more effective and safe delivery of health services (5). It is vital that there is confidence in all aspects of the quality of this information as the delivery of safe and effective healthcare depends on access to, and use of information that is accurate, valid, reliable, timely, relevant, legible and complete. Data is not only fundamental in the delivery of safe and timely care but it is essential for the appropriate care and the developmental of new services that are considered the best return for tax payer's money.

In terms of complete data availability, there are a number of major deficiencies. For example, there is very limited data available from the primary and community care sectors, from outpatients, or from emergency departments. The absence of a unique identifier for individuals across the health and social care system, and also for healthcare practitioners and organisations, results in duplication, fragmentation, increased costs and undermines the quality and safety of services. It also makes it is very difficult to follow the care pathway of an individual. (2)

In recent times with Covid 19 it has never been more important for our health system to manage data efficiently. This period in time has emphasised the importance of standardising and ensuring availability of data. We have learnt in the last two years that data is king when making decisions that impact a nation so significantly.

With the current pandemic in mind, it has shown that health information has a key role to play in healthcare planning decisions – where to locate a new service, whether or not to introduce a new national screening programme, and decisions on best value for money in health and social care provision. Although there are a number of examples of good practice, the current ICT infrastructure in health and social care in Ireland is highly fragmented with major gaps and silos of information. This results in patients and service users being asked to provide the same information on multiple occasions.

It is essential that data is accurate, relevant and timely in order to improve health and social care, to inform decision-making, monitor diseases, organise services, inform policy making, conduct high-quality research and plan for future health and social care needs, both nationally and locally. Historically in Ireland, as in many other countries, our national data collections have evolved over time in a largely uncoordinated fashion. Although there are examples of very good practice, this lack of coordination has led to a fragmented health information infrastructure.

For example, when giving a patient a drug, a nurse needs to be sure that they are administering the appropriate dose of the correct drug to the right patient and that the patient is not allergic to it. Similarly, lack of up-to-date information can lead to the unnecessary duplication of tests – if critical diagnostic results are missing or overlooked, tests have been repeated unnecessarily and, at best, appropriate treatment is delayed or at worst not given (4). Accordingly, a considerable amount of data is collected about health and social care services in Ireland. National health and social care data collections are national repositories of routinely collected health and social care data in the Republic of Ireland. They play a crucial role by providing a national overview of a particular health or social care service. There is little point in investing considerable time, effort and resources into producing a high-quality data collection if the data is not used to the maximum benefit of the population it serves. Therefore, it is essential to promote, encourage and facilitate the use of the data.

Based on international best practice, four key principles relating to health information have been identified which are based on maximising health gain for the individual and the population, specifically:
1: Health information is used to deliver and monitor safe and high-quality care for everyone
2: Health information should be of the highest quality and, where appropriate, collected as close as possible to the point of care
3: Health information should be collected once and used many times
4: Data collection should be 'fit for purpose' and cost-effective

The Role of Data Analytics in Healthcare

Data Analytics is the process of examining raw datasets to find trends, draw conclusions and identify the potential for improvement. Healthcare analytics uses the current and historical data to gain insights, macro and micro, and support decision-making at both the patient and business level.
The use of health data analytics allows for improvements to patient care, faster and more accurate diagnosis, preventative measures, more personalised treatment and more informed decision making. At the business level it can lower costs, simplify internal operations and more.
As per HIQA (2012) (3) To improve data quality you first need to measure data quality to identify what needs to be improved. Efforts to improve systems or processes must be driven by reliable data which not only allows deficiencies to be accurately identified, but also assists in prioritising quality improvement initiatives and enables objective assessment of whether change and improvement have occurred. From taking all of the above into consideration it was realised that there were silos in our data management in the Irish health care setting therefore a process was developed to improve and organise our very valuable health information. The above statement may be 10 years old but it is still as relevant as ever in 2022.

Background to the Data Specification Management Process

A National Data Dictionary and a National Release Centre for SNOMED CT have been established under the Enterprise Architecture function in the Office of the Chief Information Officer (OoCIO). Both of these functions were identified in the Knowledge and Information Plan of 2015, as being a core deliverable to realise how data as an asset could be managed.

Data quality management is a set of procedures and technologies for effectively integrating and validating data sources, securely collaborating between trusted parties, handling lifecycle systems such as aggregation and deduplication, and safely sharing the results while protecting sensitive customer information.

The importance of data quality has been highlighted in recent times when access to accurate and timely data has been needed to assist policy makers in making decisions in the management of COVID -19. A key principle in the management of all data is that it is an important asset and that it is essential to ensure that data is standardised so that it is comparable when extracted from various health information systems to inform decisions.

Currently in the HSE, there are various datasets, managed by different software suppliers, project groups and clinical groups. There is no standardised way of deciding if any one format is the correct way to identify any data element.

In 2019, the SNOMED Governance Board, chaired by HIQA made a decision to try and address this, and developed the dataset specification management process. This was based loosely on a model developed by the NHS to standardise, and manage datasets, coming through a defined process whilst having all key decision makers involved in data management in the organisation involved in the process. This proposal was submitted to Senior Management Team of the OoCIO and was approved.

A Dataset Specification Management Process (DMSP) was established and Terms of Reference agreed. This process was designed by the Business Architect (Enterprise Architecture), approved and signed off by the SNOMED Governance Board to fulfil that requirement.

The Dataset Specification Management Process brings Subject Matter Experts from core areas, within and external to the HSE, together to ensure a consistent approach to dataset specification management. Subject Matter Experts were identified by process mapping the role of all stakeholders and their role in managing datasets.

This process and group were developed to provide a forum and process to enable stakeholders who have a role in the management and analysis of data, to:
- champion their data requirements,
- contribute to an enterprise view of data
- give feedback and advice to dataset owners
- manage dataset specifications to provide foundational information governance.
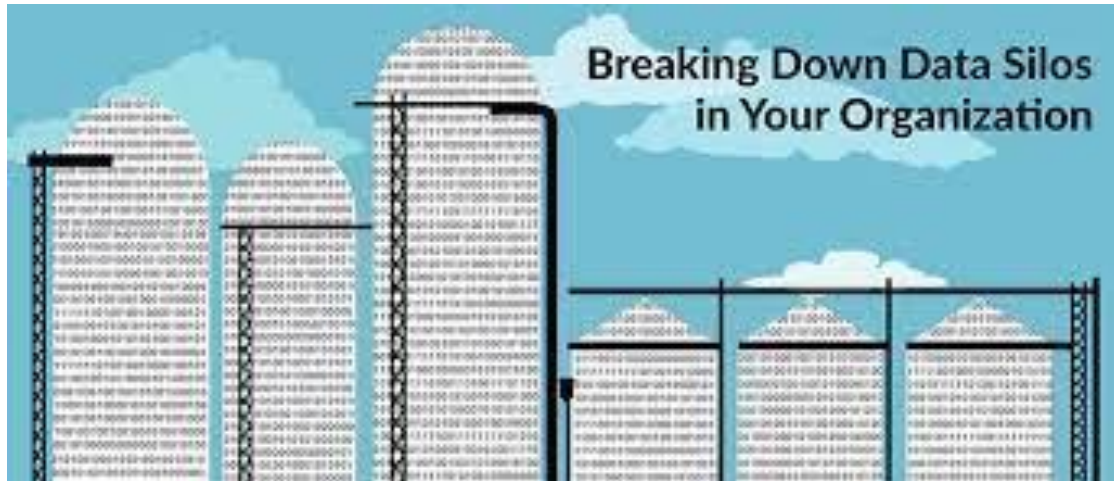
The aims of the Dataset Specification Management Process are
- a standardised uniform process that facilitates new and existing dataset specification
- a quality assurance process for new and existing dataset specifications
- the opportunity for early input from downstream functions to enhance service planning, analytics and reporting
- a reduction in re-work and costs due to errors and omissions in dataset specifications.
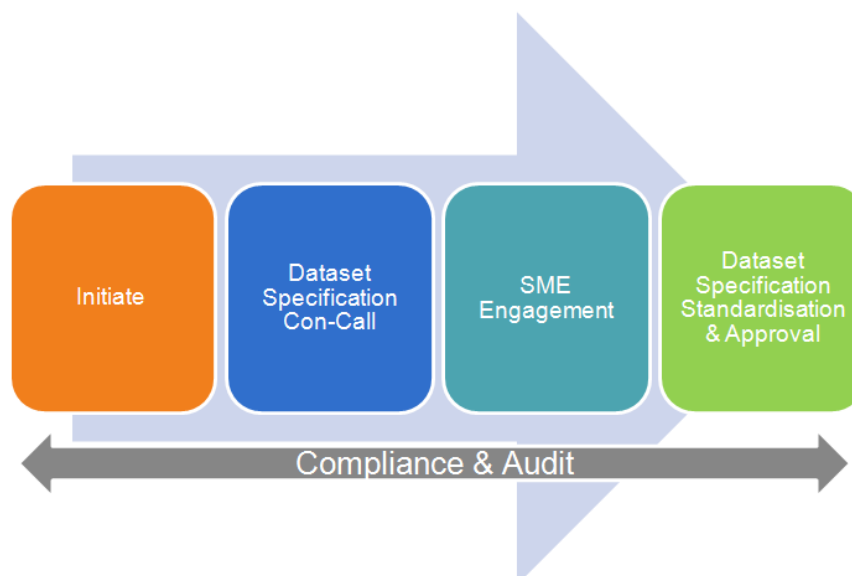
There are currently 11 Subject Matter Experts (SME) Representatives in the DSMP with 3 new members in 2021 however this is may expand in the future also. These SMEs were chosen as they are the leading experts in the HSE and National Offices for data management and expert knowledge on dataset requirements and standards.
Please review Appendix 1.0 for more details on each SME chosen for the DSMP.
Current SME representative areas:

- National Release Centre for SNOMED CT
- National Health and Social Care Data Dictionary –(NHSCDD)
- Information Governance (EA)
- Integrated Information Services (IIS)
- Health Research Board (HRB)
- Hospital In-Patient Enquiry (HIPE) – Healthcare Pricing Office (HPO)
- Business Information Unit (BIU)
- HSE Primary Care Reimbursement Service (PCRS)
- Health information and Quality Authority (HIQA) (joined 2021)
- Department of Health (DOH) (Joined 2021)
- National Office of Clinical Audit (NOCA) (Joined 2021)

The group meets monthly and the DSMP co-ordinator under the auspices of the SNOMED Governance Board and the National Release Centre for SNOMED, schedules and plans the meetings.

Ongoing Input Post Initial DSMP meeting

Each dataset requestor has an initial meeting with the DSMP and from this meeting the SMEs would then decide who would be best to help with the dataset. Depending on the size of the dataset the amount of input from the SMEs varied from a couple of meetings to more than this. Not only is it beneficial to have an SME from a particular department on board but this can open pathways to other areas of expertise to help. For example, one Dataset requestor was offered an opportunity to make a connection with the National Health and Social Care Professions Office regarding the capture of their activity. Perhaps this would not have come about without networking within the DSMP which can help with preventing any silos that unfortunately can occur within the health system.

Datasets that have been reviewed by the DSMP to date:

- **Community Health Schemes Medicine Open Data Project (CHSMODP)**
- **National Safeguarding Dataset**
- **Chronic Disease Management Programme Dataset (GP Agreement 2019)**
- **MECC Dataset (make Every Contact Count)**
- **Gynae Healthlink Discharge Summary Data Set MN-CMS**

- **National Standard on information requirements for a National Electronic Patient Summary**
- **National Psychiatric Inpatient Reporting System (NPIRS)**
- **Epilepsy EPR**
- **eServices**
- **National Nursing and Midwifery Quality Care-Metrics**
- **Covid 19**
- **Vaginal Mesh Implants & Associated Surgeries**
- **Lymphoedema Dataset**
- **Patient Flow**
- **Patient Portal Appointments Scheduling System**
- **Telehealth communication System with information sharing functionality**
- **Radiation Oncology Minimum Data Set –Phase 1**
- **ALS/MND Register and ALSCarE**

Datasets that have completed the DSMP:

Community Health Schemes Medicine Open Data Project (CHSMODP)- The
purpose of this data set is to support its use which is published monthly by the Health Service
Executive – Primary Care Eligibility and Reimbursement Service (PCERS). The dataset is based on
the claims submitted for reimbursement by Community Pharmacists in Ireland in respect of the
General Medical Services scheme (GMS), the Drugs Payment Scheme (DPS) and the Long-
Term Illness Scheme (LTI).

National Safeguarding Dataset- Incidents are recorded on a case-by-case basis currently
thus making it difficult to interconnect them. The team are working to record the 5 data fields
necessary for the Individual Health Identifier. The current system is not patient centric it is event /
issue centric. This initial dataset specification will be for a minimum dataset but overtime this will
need to be changed and extended.

Chronic Disease Management Programme Dataset (GP Agreement 2019)-
These metrics are not measures of outcomes but measures of the quality of nursing and midwifery
clinical care processes. These metrics are aligned to evidence-based standards and agreed through
national consensus in healthcare settings in Ireland. Nursing and midwifery clinical services
nationally collect indicator data monthly, and enter data into the TYC HSE system.

MECC Dataset (Making Every Contact Count)-The aim of Making Every Contact Count
is chronic disease prevention. It is about enabling health professionals to recognise the role and
opportunities that they have through their daily interactions with patients in supporting them to
make health behaviour changes. The health behaviours which are the focus of attention at the
outset are the four main lifestyle risk factors for chronic disease; tobacco use, physical inactivity,
harmful alcohol consumption and unhealthy eating**.**

Gynae Healthlink Discharge Summary Data set MN-CMS- The MN-CMS EHR System
is in use in Cork University Maternity Hospital, University Hospital Kerry, Rotunda Maternity Hospital

and the National Maternity Hospital. They wanted to develop a standard discharge summary for gynaecology patients.

Coronavirus Reference Set Ireland- Dataset for COVID 19 specifications.

Ireland Nursing and Midwifery Care Metrics Reference set- INEWS – Irish National Early Warning Score- The National Early Warning Score (NEWS) National Clinical Guideline has been updated and revised by the Irish National Early Warning System (INEWS) Guideline Development Group (GDG) under the auspices of the HSE National Deteriorating Patient Recognition and Response Improvement Programme (DPIP). The Irish National Early Warning System (INEWS) now refers to an early warning system rather than an early warning score as in the original NEWS (2013).

How to access the datasets

All of these datasets as well as being reviewed by the DSMP group have now been quality assured by the National Release Centre for SNOMED, and have been developed into reference sets that can be made available to any software vendor who will need to display, record, retrieve or use the data for analytics. This means that the clinical people who use this data every day have configured their dataset to be ready for digital health and are available to view here on the Irish SNOMED browser or here on the eHealth Ireland website under Irish refsets as a PDF.

New Dataset Potential for DSMP
- National Ambulance Service
- Breast check
- Chronic Disease management phase 2
- Dentistry CUDSH
- Orphanet and Rare Diseases.

Feedback from DSMP Clients
To date the feedback, we have received from dataset requestors has been very positive –

- "*it was very beneficial in terms of validating some of the work we have already done as well as instructing us as to what other types of data we could/should consider storing on the system. The discussion at the meeting was an excellent exercise for the project team in terms of reflecting on the purpose of the system and that the data we are collecting was necessary and sufficient for that purpose.*"

- *"yes the process did force us to refine our terminology and clarify it"*.

- *"Everyone was extremely helpful when we got them engaged on who we were and what we were trying to achieve.*
  I think there was value in us doing this work as "Domain Experts" and value for data dictionary process also with the submission of a new dataset.

-

- *'We got a great learning on all the nuts and bolts of data standardisation and how different aspects of data relate to each other'*

Overall improvements required from previous DSMP clients

The feedback we have received has been very positive but as this is a new process there is always an opportunity to improve our service. By having a bit more clarity on the process of the DSMP for data requestors this would improve the datasets presented in a less raw state before approaching the DSMP. Dataset Owner / Requestors are having some difficulties with applying / getting approval for changes identified as necessary from the DSMP.

Recent review of DSMP with Subject Matter Experts in October 2021

- Increase number of SME's to DSMP group - HIQA, DOH and NOCA recently joined
- Regular review of service provided by DSMP
- Group plan to continue with monthly meetings and develop a prioritisation process
- Terms of Reference reviewed and updated.

Benefits of the DSMP to date

- Access to the leading authorities in data management in the HSE, DOH, HIQA and other national bodies.
- Noted reduction in duplication of datasets due to an overall awareness nationally of data management hence further reducing silos in the organisation.
- There is a defined process for managing datasets
- Built a network of authorities involved in managing and retrieving data in the HSE to work towards a common goal.

Future Plans

The DSMP follows a set of Enterprise Architecture Principles that are integral in the development of datasets not only for here and now but future proofing for interoperability and ensuring that the data standards are set to support the delivery of integrated healthcare.

These principles are:

- Data is an asset that has value to the enterprise and is managed accordingly.
- Users have access to data necessary to perform their duties; therefore, data is shared across enterprise functions and organisations in accordance with legislation.
- Data should be accessible for all users to perform their work. Users should have access to local and core national records relevant to their role and within legislative parameters.
- Data is defined consistently throughout the enterprise, and the definitions are understandable and available to all users.
- Each data element has a trustee accountable for data quality.
- Semantic interoperability is achieved through the use of SNOMED CT in all clinical information systems.
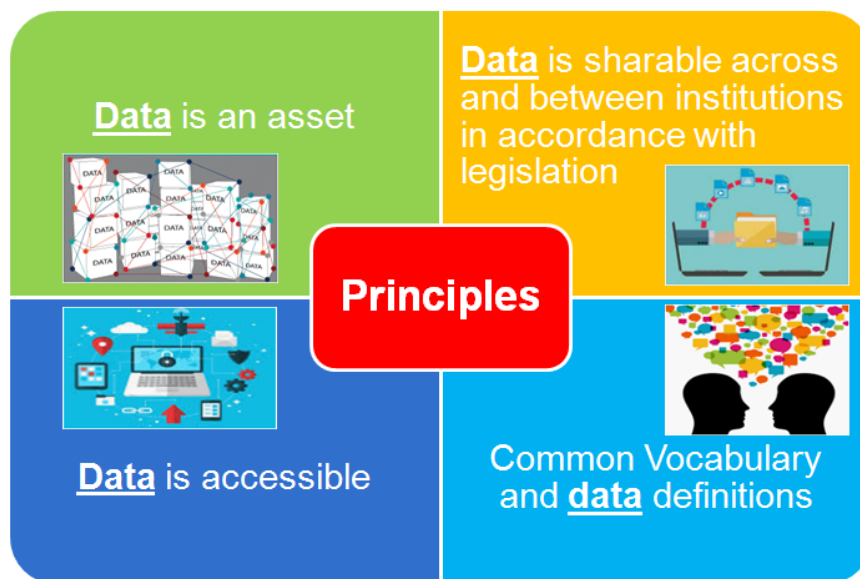
Enterprise Architecture Principles

As outlined in the HIQA report published October 2021, there is an urgent need for further emphasis on the way in which healthcare information is managed, governed and regulated. A process such as the DSMP will streamline all datasets in Ireland, ensure standardisation, avoid duplication of resources and essentially provide a high standard of care by integrating care. For further information on this report please see link in **Appendix 1.2**. The DSMP continue to grow this process for all datasets in the organisation to improve standards throughout our data management within the Health System.

For any further information please contact:

Theresa Barry **Clinical Terminology Architecture Lead** Theresa.Barry@hse.ie

Yvonne Coughlan **Business Analyst National Release Centre for SNOMED CT yvonne.coughlan2@hse.ie**

Appendix:

Appendix 1.0

Name of Subject Matter Expert	Reason for being chosen for DSMP
SNOMED CT	National Release Centre for SNOMED CT-Standardising Terminology for EHR Systems
NHSCDD-HSE National Health and Social Care Data Dictionary	Data Dictionaries for data are useful to describe the 'rules' which data in a particular system need to abide by. For any new projects or applications being introduced into the HSE, the HSE NHSCDD can help with the creation of the datasets, alignment of existing datasets and assist vendors through the development of evolving datasets specifications.
Information Governance (EA)	Information governance provide standards and guidelines on the handling of data and giving the tools to organisations need to achieve the standards required for information such as that contained in a healthcare record is handled legally, securely, efficiently and effectively in order to deliver the best possible care to people who use health and social care services.
Integrated Information Services (IIS)	The IIS is a centralised service that creates and maintains policies, standards and governance around the HSE's data. It is the main Data Analytics Service for the wider Health Service Executive.
Health Research Board (HRB)	The Health Research Board (HRB) is a State Agency under the Department of Health. It is responsible for supporting and funding health research, generating health information and promoting the use of evidence in policy and practice. The HRB's National Health Information Systems Unit manages four national health information systems on behalf of the Department of Health. These systems 1) produce data and information to assist the Department with policy and planning 2) enable the HSE to report on its service planning processes and 3) generate research that is relevant to both policy and practice.
Hospital In-Patient Enquiry (HIPE)- Healthcare Pricing Office (HPO)	HIPE Scheme is a health information system designed to collect demographic, clinical and administrative data on discharges from, and deaths in, acute public hospitals nationally. HPO is to oversee all functions associated with the operation of the HIPE database, including the development and support of the data collection and reporting software, training of coders, data quality, audit, data analysis and reporting, and responding to requests for information.
Business Information Unit (BIU)	The HSE's Business Information Unit (BIU) gathers and maintains a central repository of service data from all hospital and community based health services nationwide. Extensive amounts of data are collected, collated and checked within this Unit. This data, is used as part of the National Performance Oversight process and in other levels of performance management as part of the Performance Accountability Framework
HSE Primary Care Reimbursement Service (PCRS)	The PCRS compiles statistics and trend analyses which are provided to other areas within the HSE, the Government, customers, stakeholders and to members of the public.

Health Information and Quality Authority (HIQA)	HIQA Working in conjunction with a wide range of stakeholders, we aim to improve the quality and safety of health and social care services by setting national standards and publishing guidance.
Department of Health (DOH)	The Department of Health is responsible for the overall strategy and planning of National service plan and SlainteCare.
National Office of Clinical Audit (NOCA)	NOCA manages national clinical audits that aim to improve patient care and outcomes. NOCA's findings enable the healthcare system to act to improve care where standards are not followed. NOCA enables the Irish healthcare system to continually improve its standards of care via maintenance of a portfolio of prioritized national clinical audits, standardised against national and international criteria. Not only do NOCA enhance accessibility to validated data for persons who use, manage and deliver healthcare, our clinical audits help to improve patient outcomes, and create positive change locally and nationally.
Enterprise Architecture Office of the CIO, Eamon Coyne	Enterprise Architecture is responsible for supporting the strategic development of technical architecture, technology and operational capabilities in a standardised and aligned manner.

Appendix 1.2:

Follow this link: https://www.hiqa.ie/reports-and-publications/health-information/need-reform-irelands-national-health-information-system

References:

1. HIQA (2011) Recommendations for Unique Health Identifiers for Healthcare Practitioners and Organisations {Accessed December 18th 2021} https://www.hiqa.ie/sites/default/files/2017-01/Health-IDs-Practitioners-Organisations-Summary.pdf

2. HIQA Guiding Principles Data Collections (2013) { Accessed December 8th 2021} https://www.hiqa.ie/sites/default/files/2017-01/Guiding-Principles-Data-Collections.pdf

3. HIQA Data Quality Guide (2012) {Accessed December 18th 2021} https://www.hiqa.ie/sites/default/files/2017-01/Data-Quality-Guide

4. HIQA Healthcare Interoperability Standards (2013)) { Accessed December 8th 2021} https://www.hiqa.ie/sites/default/files/2017-01/Healthcare-Interoperability-Standards.pdf

5. Department of Health (2019) Sláintecare Implementation Strategy {Accessed January 14th 2022} https://www.gov.ie/en/publication/0d2d60-slaintecare-publications/#2019-action-plan