Common Data Space for Ireland
IRL-DataSpace

IRL-DataSpace & IRL-DSSC Position Paper

A Common Data Space & Support Centre for Ireland

Venkatesh Kannan

Niall Wilson

Jean-Christophe Desplat

Eoin Kenny

Jenny O'Neill

Roberto Sabatino

Glenn Wearen

ICHEC Irish Centre for High-end Computing

HEAnet Ireland's National Education & Research Network

Common Data Space for Ireland
IRL-DataSpace

Contents

Keywords

Data Space, Data Service, Support Centre, Research, Innovation, Ireland

Suggested Citation

"IRL-DataSpace Position Paper: Building a Common Data Space for Ireland, May 2024"

Common Data Space for Ireland
IRL-DataSpace

Executive Summary

This position paper sets the context and recommends a framework to build a Common Data Space for Ireland (IRL-DataSpace) to address data-driven global challenges through collaborative Research and innovation (R&I) between academia, industry and the public sector. The purpose of IRL-DataSpace is to cater to R&I-focused participation and usage of the data ecosystem by academic, public sector and industry participants for dealing with grand challenges and informing better public policy-making. The usage of IRL-DataSpace is not intended for commercial services.

This position paper aligns with the objectives and directions laid out in Impact 2030: Ireland's Research and Innovation Strategy[1] and Open Data Strategy 2023-2027[2], and also aligns with related strategies/initiatives including the National Action Plan for Open Research 2022-2030[3], AI - Here for Good: National Artificial Intelligence Strategy for Ireland[4], Climate Action Plan for Ireland[5], National Smart Specialisation Strategy for Innovation 2022-2027[6], and Common European Data Spaces[7] (part of the European Strategy for Data[8]).

The proposed framework and roadmap will deliver IRL-DataSpace by accelerating and supporting a structured synergised unification of the national data ecosystem (composed of academic and public research bodies, and enterprise organisations where relevant) which is currently disparate and fragmented. IRL-DataSpace will be a federation of data stakeholders, data systems and services at a national level that also links with relevant European resources. The beneficiary activities and exploitation of IRL-DataSpace target multidisciplinary and interdisciplinary R&I collaborations driven by academic and public bodies with industry partners as required.

This position paper is led by HEAnet (Ireland's national education and research network provider) and ICHEC (Ireland's national centre for high-performance computing and data services). The shared vision between HEAnet and ICHEC is to expedite the establishment of a national-level Data Space Support Centre for Ireland (IRL-DSSC) through which to engage stakeholders across the community for their participation and co-development. The IRL-DSSC will play a pivotal role in advancing the current national data ecosystem into a federated, FAIR (Findable, Accessible, Interoperable, Reusable), high-SLA and cost-effective ecosystem with robust digital platform resources, services for data sharing and governance, and a national-level professional support team. The IRL-DSSC will engage with and leverage expertise from the EU Data Spaces Support Centre[9].

ICHEC and HEAnet are proposed to co-lead IRL-DSSC with sponsorship from relevant government departments and participation of relevant stakeholders. ICHEC and HEAnet have available the expertise that can be leveraged for this as the national providers of digital infrastructures and services across computing, data management and networking technologies. Together, IRL-DataSpace and IRL-DSSC will deliver R&I impact across many disciplines in environment, society, economy at national and international levels, by dealing with grand challenges and informing better public policy-making as outlined in the Impact 2030 Strategy.

Complementing this paper on IRL-DataSpace, the "Vision to Implement the European Open Science Cloud (EOSC) in Ireland" is outlined by HEAnet in consultation with ICHEC and the consortium of Irish Research Libraries (IReL). These two papers are aligned and linked to enable the connection of Data Spaces to EOSC.

---

[1] https://www.gov.ie/en/publication/27c78-impact-2030-irelands-new-research-and-innovation-strategy/

[2] https://www.gov.ie/en/policy-information/8587b0-open-data/#open-data-strategy

[3] https://doi.org/10.7486/DRI.ff36jz222

[4] https://enterprise.gov.ie/en/publications/national-ai-strategy.html

[5] https://www.gov.ie/en/publication/79659-climate-action-plan-2024/

[6] https://www.gov.ie/en/publication/4b219-national-smart-specialisation-strategy-for-innovation-2022-2027/

[7] https://digital-strategy.ec.europa.eu/en/policies/data-spaces

[8] https://digital-strategy.ec.europa.eu/en/policies/strategy-data

[9] https://dssc.eu/

Common Data Space for Ireland
IRL-DataSpace

Glossary of terms

Terms	Description
Data management	A combination of practice and technologies for systematically collecting, storing, sharing and governing the use of data.
R&I activity	Research and innovation work conducted by an institution, research centre, project or researcher.
Institutions, research centres	Entities involved in R&I activities including, but not limited to, HEIs, SFI research centres, EI technology centres, State Agencies and Departments.
Data	A collection of datasets, software tools, documents and metadata describing the data.
Data ontology	A methodology and model for organising and linking structured and unstructured data from various sources and formats using entities, properties and relationships between them.
Database	An organised electronic collection of structured data.
Data system	A data repository, or data warehouse, or data lake, or data space.
Data repository	A storage space in which to deposit data.
Data warehouse	A centralised data repository in which to integrate data from one or more disparate sources for targeted analysis.
Data Lake	A digital infrastructure system for efficient and secure storage, management, processing and service.
Data Space	A data ecosystem built around commonly agreed building blocks, enabling an effective and trusted sharing of data among participants who may be data systems, their providers and users.
EU DSSC	EU-funded Data Space Support Centre to promote, coordinate and support the deployment of common data spaces, and reuse of technologies, processes, standards and tools across sectors and dataspace implemtation activities.
IRL-DataSpace	A national-level fedetated system of systems composed of data stakeholders, data systems and services that also links with relevant European and global resources.
IRL-DSSC	Ireland-focused Data Space Support Centre providing trusted neutral national-level technical coordination and support for the national R&I stakeholders, and bridging national and international actors and beneficiaries for effective implementation, operation and sustenance of IRL-DataSpace.

List of figures

List of tables

Common Data Space for Ireland
IRL-DataSpace

1. Context

A robust, federated and efficient national data ecosystem is essential to pursue, accelerate and enable the multi- and inter-disciplinary as well as inter-institutional R&I activities that target global challenges across different disciplines – societal, environmental and economic – to inform better public policy-making. These R&I activities are increasingly data-driven and linked to policy-making. Hence, the data ecosystem is critical to achieve the objectives and relevant impacts envisioned in key national and relevant European strategies:

Impact 2030: Ireland's Research and Innovation Strategy

Effectively harness data from research, public services and enterprises through improved data sharing and national-level data infrastructures to deliver on wider national and European agendas and R&I for public policy development.

Open Data Strategy 2023-2027

Accelerate digital transformation through the Open Data Strategy Framework in Action through R&I for evidence-based decision making and connect communities across sectors and institutions through a trusted FAIR and open data ecosystem.

National Action Plan for Open Research 2020-2030

Establish a mature ecosystem of infrastructures to support responsible management and sharing of research data and other outputs in line with FAIR Data principles.

National AI Strategy for Ireland

Build a trustworthy data governance framework with quality high-integrity datasets, with access protocols and services to process data with governance and privacy, spanning a federation of data management and HPC infrastructure for successful exploitation of AI.

Climate Action Plan for Ireland

Understand the causes and monitor the impacts of climate change and the measures to address it through cross-institutional data platforms and services to collect, generate, store and manage national databases and open access data products.

National Smart Specialisation Strategy for Innovation

Digital transformation to create new and modify existing business processes, culture and customer experience through use of digital and data technologies to deliver scale, productivity and innovation advancements.

European data strategy

The European data strategy aims to make the EU a leader in a data-driven society by creating a single market for data allowing it to flow freely within the EU and across sectors for the benefit of businesses, researchers and public administrations. Supporting European activities include:

> ***European Data Act***
> Entered into force in January 2024, the EC aims to make more data available for use via new rules on who can use and access data and for which purposes across all economic sectors in the EU.

> ***European Data Governance Act***
> A key pillar of the European strategy for data, the Data Governance Act seeks to increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data. Directive (EU) 2019/1024 on open data and the reuse of public-sector information.

> ***Common European Data Spaces***
> Unleash the potential of data-driven innovation in a trustworthy and secure manner by EU businesses, public administrations and researchers through the development of new data-driven products and services, forming the core of an interconnected and competitive European data economy.

In this context, it is essential to acknowledge the following aspects:

- **Data has a lifecycle**: Data in the research, public sector and enterprise communities has a lifecycle where it progresses through various stages of maturity, quality, publication-readiness and preservation. It is essential for institutions, sectoral communities and nationalf-level organisations to coordinate, robustly manage and govern datasets in the context of this lifecycle to facilitate open and FAIR Data. This is elaborated in Section 2.1.

- **Data systems are complex and varied**: The ecosystem of data systems in Ireland, to support Open Research and FAIR Data through different stages of their lifecycle, is complex and varied. It encompasses data storage infrastructures, discipline-specific/multi-disciplinary repositories, institutional data repositories, networking and computing infrastructures, meta-infrastructures (for cross-cutting services such as Authentication and Authorization Infrastructure (AAI), persistent identifiers (PIDs), catalogues, etc.), and varied policies/practices for data management. Also, the terms data repository, warehouse, lake and space are often used inaccurately and interchangeably leading to confusions and misinformation. While some disciplines and institutions (including research centres, agencies, etc.) have well-defined standards and protocols for data management, others do not have specifications for management, sharing and governance of data. This is discussed in Section 2.2.
- **Data ecosystem is a federation**: The ecosystem of data systems provided/operated by different stakeholders will need to be unified as a federated system of systems through robust and efficient policies, processes and technologies based on the FAIR Data principles. There cannot be a centralised solution within a single storage infrastructure (thematic or national) due to inherent data ownership, governance, sensitivity and security requirements.
Cross-cutting and mission-critical technical services will have to be operated and delivered by national-level providers (such as HEAnet and ICHEC) in close partnership with members in the federated system of systems. An approach for Ireland to achieve this is outlined in Section 3.
- **Data management standards and practices vary across domains**: There are several frameworks and supporting technologies to implement and govern a federated data ecosystem. These include domain-agnostic approaches such as Common European Data Spaces, European Open Science Cloud (EOSC), European Collaborative Data Infrastructure (EUDAT), as well as thematic solutions such as Earth Systems Grid Federation (ESGF), European Geological Data Infrastructure (EGDI), Health Data Access Sharing Storage Linkage (DASSL), Earth Observation Exploitation Platform Common Architecture (ESA EO EPCA), European Plat Observing System (EPOS), and many more. A list of exemplars is outlined in Section 4.

In order to deliver the impacts envisioned by the aforementioned strategies, and bridge the gaps and challenges outlined in the above-mentioned aspects, it is essential that Ireland pursues the following:

- **Map the Irish data ecosystem**: A comprehensive and consistent mapping of the data ecosystem and data lifecycle in Ireland is essential by working closely with the key stakeholders (sectoral, thematic, platform and service providers, end-users in academia, public-sector and industry, and the Impact 2030 Advisory Forum and/or Steering Group). The focus should be on enabling R&I activities that address national and global challenges (societal, environmental and economic) that require muti- and inter-disciplinary as well as inter-institutional and cross-border collaborations.
- **Develop a Common Data Space for Ireland**: Coordinate the development of a synergised and efficient federated national ecosystem of data systems and services for seamless FAIR Data, secure, human and machine-level sharing, linking and governance of datasets and associated products, and their interface with institutional and/or public compute platforms. This federated national data ecosystem would be the Common Data Space for Ireland (IRL-DataSpace).
- **Establish national-level technical coordination and support**: To achieve a functioning federated system of systems – the Common Data Space for Ireland (IRL-DataSpace) – a national-level technical body is essential to coordinate, support and provide technical expertise to the different stakeholders in the federated data ecosystem. This body will coordinate, support and synergise the different data holders and governors, data system operators, policy-makers and end-users. This would be through establishment of a nation-level Data Space Support Centre for Ireland (IRL-DSSC) as an operational body providing professional and technical expertise, coordination and support to national stakeholders based on sectoral, national and European standards.

**IRL-DataSpace and IRL-DSSC are an essential component for Ireland
to deliver these key national and European strategies.**

2. Data lifecycle and ecosystem

As outlined earlier, the data ecosystem is complex and varied, primarily due to the heterogeneity of the data sources, providers, data systems, maturity of the data services, users across different sectors and domains, as well as specific requirements for data governance, security and sensitivity. Most importantly, due to the nature of the targeted activities and impacts for our national and global challenges, it is necessary to scale up and accelerate data-driven multi- and inter-disciplinary R&I in the context of cross-sectoral national- and European-level coordination through a federated system of data management systems and services.

In order to highlight the most significant aspects to illustrate the complexity and heterogeneity involved,

- Section 2.1 outlines a high-level overview of data lifecycle that is most relevant for R&I activities in collaborations among academia, public sector and relevant enterprises.
- Section 2.2 presents a representative mapping of the observed ecosystem of data systems, services and stakeholders at the national and relevant regional levels.
- Section 3 outlines the first steps for an approach to synergise and unify the existing data-related ecosystem stakeholders and resources, along with proposed requirements and gaps to be addressed to achieve a synergised Common Data Space for Ireland (IRL-DataSpace).
- Section 4 summarises the relevant data management frameworks and activities that should be leveraged for coordinating, synergising and developing the IRL-DataSpace.

2.1. Data sources and lifecycle

Data is collected and generated from a wide range of sources for R&I activities as well as decision-/policy-making purposes. Such data sources may be grouped into four main types as illustrated in Figure 1 – observational, experimental, simulation, and derived. The importance in understanding and embedding these types of sources for the Common Data Space for Ireland (IRL-DataSpace) is that the mode of collection and generation may affect the data management policies and practices. For instance, if data from different sources and/or providers need to be combined, then sharing policies, standards for formats and metadata, access interfaces and security consideration are key to lower the barriers for institutions/activities/projects to share and link data. Another example is when data is hard or impossible to replace and is of mission-critical nature at national/institutional levels, then quality of service, service level agreements, redundancy to reduce loss, sustainable funding and operation of the data systems are essential priorities to address.
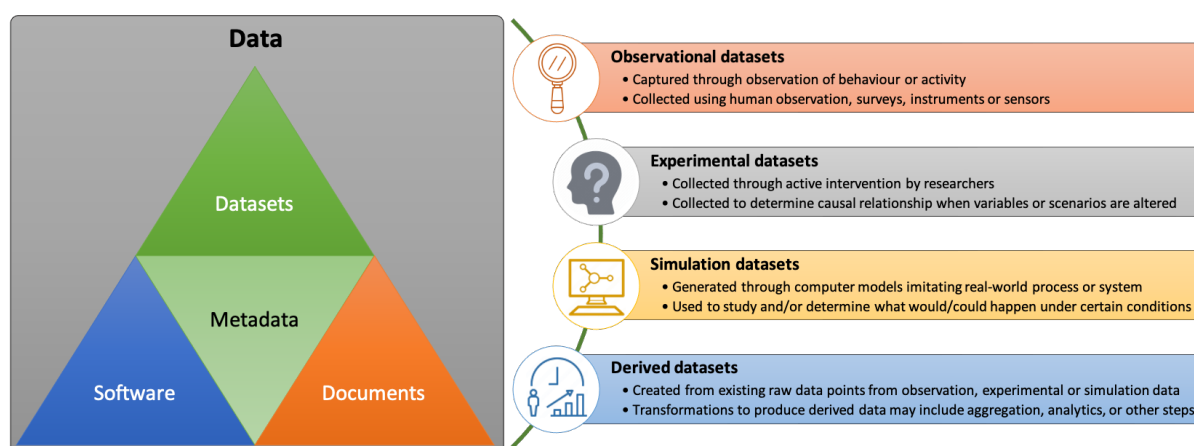
Figure 1: Types of data sources relevant to research and innovation

One of the objectives of Ireland's National Action Plan for Open Research (2022-2030) is enabling FAIR Data and outputs from R&I activities. On one hand, it is important to establish that the Irish landscape of data ecosystem is complex and varied (elaborated in Section 2.2). On the other hand, all types of data from the varied sources have a management lifecycle as outlined in Figure 2, and all R&I-relevant data need an enabling mechanism to transition into FAIR Data, which currently is not prevalent or universal.

Common Data Space for Ireland
IRL-DataSpace

Figure 2: Research data management lifecycle[10]

This enabling mechanism includes services at a national-level that provide tools, best practices and services (technical and mentorship) to support effective management and maturation of the data to be FAIR for suitable sharing at each stage.
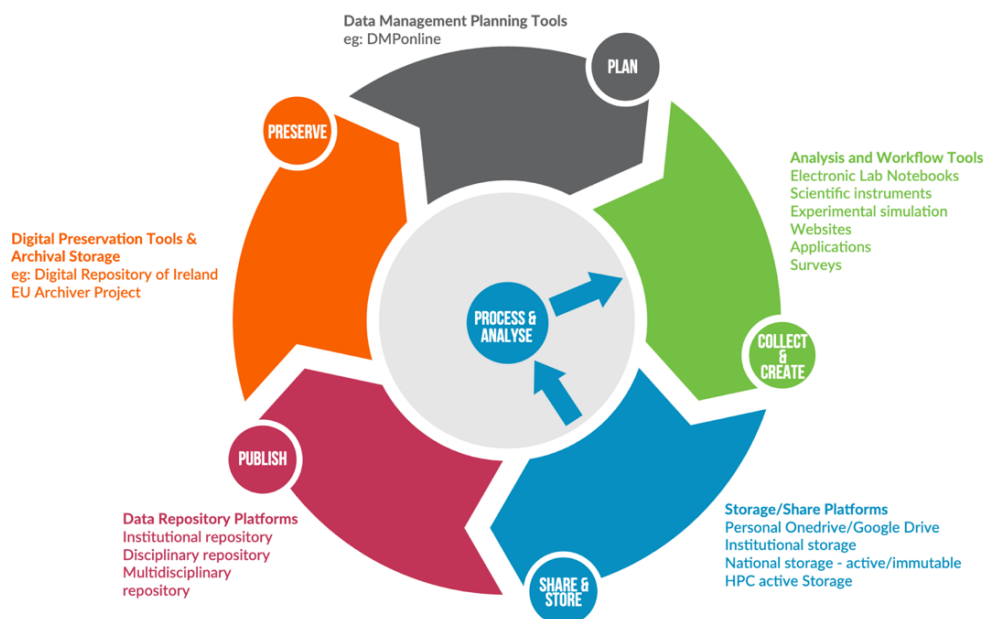
> **Data has a lifecycle**
>
> - Data in the research, public sector and enterprise communities has a lifecycle. It progresses through various stages of maturity, quality, publication-readiness and preservation.
> - It is essential that institutions, sectoral communities and national-level organisations coordinate, robustly manage and govern datasets to mature them to open and FAIR Data.

2.2. Outline of R&I-relevant data ecosystem in Ireland

We define an R&I activity as one that is conducted by an institution, research centre, project or researcher. Within any R&I activity, each dataset from different types of sources during its management lifecycle (outlined in Section 2.1) undergoes a complementary lifecycle based on its maturity and the activities performed on it by its stakeholders (owners, providers, users). Based on observations gathered from key R&I domains in Ireland (including but not limited to Earth Sciences, Environmental Sciences, Life Sciences, Urban Sciences, Material Sciences, and ICT), we categorise a dataset's lifecycle as outlined in Figure 3.
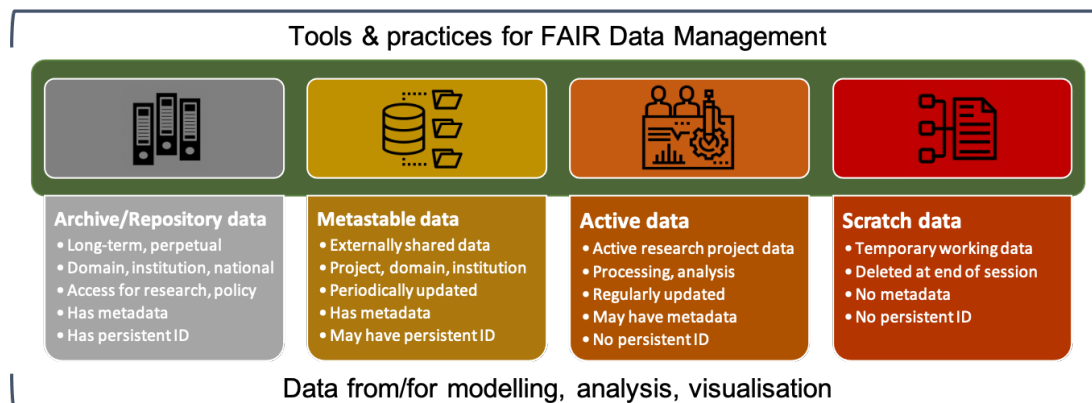

Figure 3: Data lifecycle based on maturity and activity

---

[10] National Action Plan for Open Research, 2022-2030

A dataset in any one of these four states may transit to any other state through an action (by human or machine) of making a distinct copy. While the original remains in the current state, the copy exists in the new state with or without a persistent identifier (PID) depending upon the purpose of its use.

- **Scratch data**
  - This is temporary data on which a specific work or task may be carried out by an R&I activity, and by default will be deleted at the end of a work session. Explicit effort is required to save a copy of this data for later access or sharing.
  - For instance, these may be data that are analysed or processed in the context of a computational model or analysis application.
  - No metadata are created/available for this dataset, and no persistent identifier (PID) is associated with it for unique identification. Its maturity is very low w.r.t. FAIR-ness.

- **Active data**
  - This is data exists in a R&I activity during an Active Research Phase(s) (e.g., sub-projects, work packages, tasks) which may involve conduct of quantitative/qualitative analysis, computations, modelling, visualisation or other such tasks. This data is actively updated on a regular basis, most likely, based on the results from scratch data.
  - Owners of the active data share it with other members for collaborative work or tasks within the same R&I activity that may span different institutions and/or countries.
  - Generators of the active data may optionally create and attach metadata when storing and sharing based on global, domain or individual standards. No persistent identifier (PID) is associated with active datasets. Its maturity is moderate w.r.t. FAIR-ness.

- **Metastable data**
  - This data is hosted in an R&I activity level to be retained for longer periods of time, is expected to be updated and re-used periodically in a systematic way as a part of future Active Research Phase(s) and/or shared with other R&I activities without significant efforts and quality towards curation.
  - Metadata will be created and attached to these datasets based on global or domain-specific standards. A persistent identifier (PID) may optionally be created and associated with the dataset for unique identification. Its maturity is expected to be moderate to high w.r.t. FAIR-ness due to the intent to publish or share the datasets with collaborators within and/or outside the R&I activity.

- **Archive/Repository data**
  - This data is meant to be for long-term preservation and perpetual availability at domain, institutional, national or global levels. The availability and sustainability are critical for R&I and/or decision-/policy-making purposes in academic, public sector and/or industry.
  - All datasets should have associated metadata and persistent identifier (PID) published. Consequently, they are high w.r.t. FAIR-ness.
  - Due to the high availability and sustainability requirements, the data systems on which archive data are managed require high quality of service, service level agreement, and long-term funding and operational requirements.

Based on engagements with a broad range of R&I activities (institutions, research centres, projects, research groups, independent researchers) across disciplines at national and European levels, it is observed that data systems currently exist in two flavours – basic and mature – to manage and share data (in the context outlined in Figure 1, Figure 2 and Figure 3). These two flavours are illustrated in Figure 4:

- **Basic Data System**
  - These data systems at an R&I activity level are typically a collection of data repositories used within the R&I activity and its personnel. Depending on the R&I activity, these data repositories and their storage systems may be independent, siloed from each other or well-structured and interconnected with each other at the R&I activity level.

- No access gateway (e.g., through user/programming interfaces or software libraries) or catalogue of data resources/services are available on such basic data system.
- Any access to the datasets by internal or external staff (who are not data owners) requires direct requests to be sent to the data owners/providers. Ingestion of data into basic data system, export of data from the basic data system is primarily on a one-to-one basis.
- Basic Data Systems have very low ability to enable FAIR-ness of its data. Consequently, this causes major challenges and limitations, including duplicated/wasted costs and efforts, in manually sharing and inefficiently moving data.
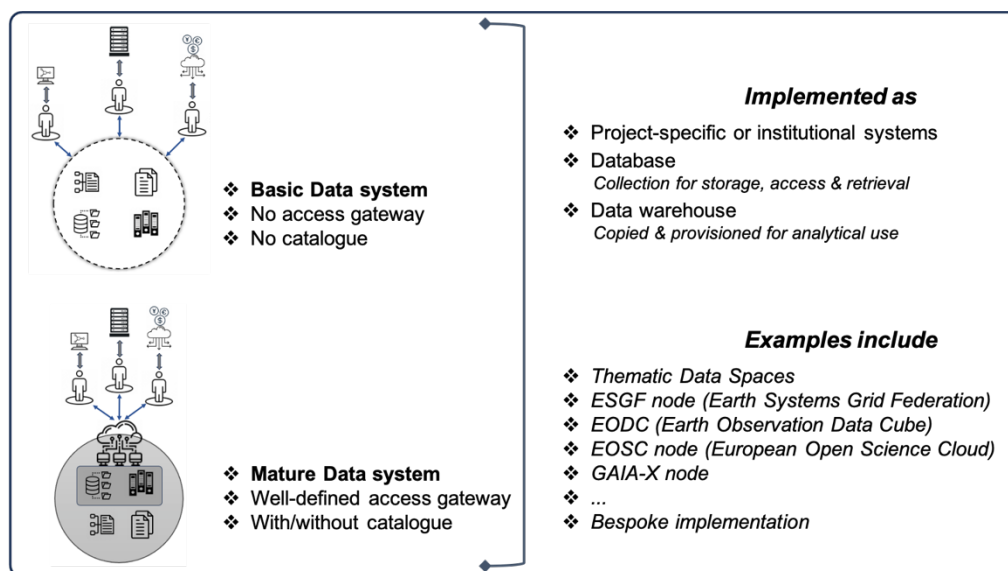

Implemented as

❖ Project-specific or institutional systems
❖ Database
  Collection for storage, access & retrieval
❖ Data warehouse
  Copied & provisioned for analytical use

Examples include

❖ *Thematic Data Spaces*
❖ *ESGF node (Earth Systems Grid Federation)*
❖ *EODC (Earth Observation Data Cube)*
❖ *EOSC node (European Open Science Cloud)*
❖ *GAIA-X node*
❖ *...*
❖ *Bespoke implementation*

❖ **Basic Data system**
❖ No access gateway
❖ No catalogue

❖ **Mature Data system**
❖ Well-defined access gateway
❖ With/without catalogue

Figure 4: Ecosystem of R&I-relevant data systems in Ireland

- **Mature Data System**
    - These data systems at an R&I activity level are well architected, engineered and operated with well interconnected component data repositories and storage systems.
    - Typically, a catalogue of data resources and data-related services is published for all users of the mature data system, along with clear processes and tools for implementing data management and governance policies.
    - A mature data system also provides an access gateway which may be composed of user/programming interfaces, software libraries, or graphical interfaces). This enables integrating the mature data system in a federation of data systems in national, regional and global ecosystems.
    - Mature data system enables a high amount of FAIR-ness of its data. This alleviates part of wasted costs and efforts that otherwise would have been spent in inefficient duplication and movement of data through redundant manual mechanisms.

In basic and mature data systems, users that gain access to the data typically copy/move the required data to local/external systems for conduct of R&I tasks (e.g., processing, modelling, analysis, visualisation). The local/external systems may be personal workstations, institutional computing platforms or commercial cloud services. It is important to note that this is a non-trivial and costed step because data has gravity when moved, policies related to its locality and governance apply, and proximity between storage and computing environments impacts efficiency of the R&I tasks.

It is important to note that two aspects about the data systems (basic or mature):

- Data systems catering to R&I activities are typically implemented for specific projects or institutions for a fixed-term which often leads to challenges related to sustainability of the data availability and operating the data services beyond the funding grants. This brings major risks, particularly to data systems that serve Archive Data.

- In many scenarios, such data systems are implemented as databases and data warehouses with different underlying infrastructure for storage. The approaches to architect, engineer and operate such data systems (basic or mature) may be based on mature specifications/frameworks (e.g., Earth Systems Grid Federation, Open Data Cube for Earth Observation and Geospatial data, European Open Science Cloud, GAIA-X, etc.) or bespoke architectures drafted by individual institutions or projects.

Furthermore, the implementation and operation of a data system (basic or mature) will be typically composed of some/all of the components illustrated in Figure 5. These include hardware infrastructure, software, datasets and personnel (technologists that build/operate the data system and data governors).
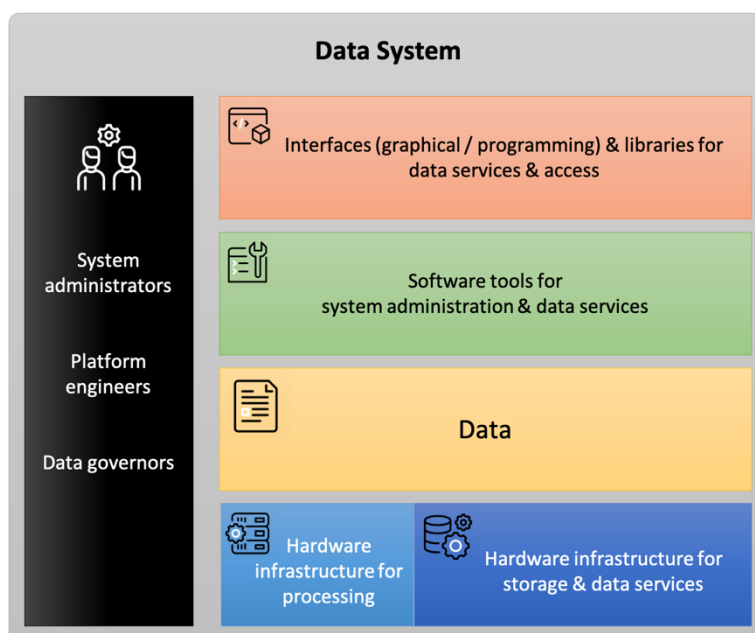

Figure 5: Technology & personnel components of data systems

It is important to note that Ireland's Open Data Portal[12] promotes innovation and transparency at Irish Public Sector organisations through publication of their open data. This is a portal provides cataloguing and searching thereby enabling Findability and Accessibility (F and A from the FAIR Data principles) of the published datasets. While many R&I activities from public sector organisations list their datasets on this portal, there remains the gap to build a national-level Common Data Space for Ireland that enables all R&I-focused data systems to participate, register and ensure full FAIR-ness and enable collaboration through a rich pool of technical tools and support across academia, public sector and industry.

> **Data ecosystem in Ireland is complex & varied**
>
> - Some institutions/domains have mature data systems (enabling high FAIR-ness), while others have basic data systems (low ability to enable FAIR-ness).
> - Ireland's ecosystem of diverse data systems will need to be unified as a federated system of systems based on FAIR principles, practices and technologies.
> - Data sharing cannot be through a centralised solution within a single storage infrastructure due to inherent requirements of data ownership, governance, sensitivity and security.
> - Data has gravity. Consequently, copying/moving data between different systems for inter-/multi-disciplinary and/or cross-institutional R&I is expensive, inefficient and not scalable.

---

[12] Ireland's Open Data Portal https://data.gov.ie

3. Approach to synergise R&I-relevant data ecosystem in Ireland

Sections 2.1 and 2.2 outline the currently varied, complex, disparate, fragmented and often siloed data systems in the Irish ecosystem. The status quo imparts a high entry barrier, significantly reduces the efficiency of researchers, and inherently limits multi-/inter-disciplinary and cross-institutional collaborative R&I among academia, public sector and industry. This is due to the ineffective use and limitations in extracting value from the data (open and closed) from different stakeholders through technical systems, tools and practices that are truly FAIR-compliant.

This limitation and challenge have to be overcome collectively at a national-level especially when working towards the global challenges (societal, environmental and economic) that are prioritised in key national and European strategies (outlined in Section 1) that rely on multi-/inter-disciplinary and cross-institutional collaborative R&I among academia, public sector and industry.

To overcome this limitation and challenge, it is proposed to synergise and unify the existing data-related ecosystem stakeholders and resources through the creation of IRL-DataSpace (the Common Data Space for Ireland). Key benefits and value of creating IRL-DataSpace include the following:

- **For data system operators**: Collectively increase, accelerate and align the maturity and FAIR-ness of data systems across Ireland that are relevant and essential for R&I activities at institutional, domain, national and regional levels.
- **For data systems**: Implement standardised digital interfaces, processes and practices across all mature data systems to ensure interoperability across all their technical and personnel components (Figure 5). Breakdown current siloes between R&I-related data systems, institutions and activities through well-defined interoperable data sharing policies, practices and tools.
- **For data holders and governors**: Empower data owners, providers and governors to make their data available for seamless access and reuse based on consistently implemented FAIR technologies and processes for R&I activities.
- **For R&I activities**: Provide a single point of entry for all R&I-focused users, projects and activities to access R&I-relevant data (open and closed) that their owners/providers share at national and global levels. Ensure fair, transparent, proportionate, non-discriminatory, secure and privacy-preserving access using well-defined and trustworthy technical and data governance mechanisms.
- **For funders and policy-makers**: Establish a synergised, well-connected and coherent national-level data marketplace that respects national and EU rules and values, removes current technical and human barriers/inefficiencies in sharing and linking data across domains and institutions, and delivers maximal impact and return on investment in data-dependent R&I activities, systems, infrastructure and personnel.

3.1. IRL-DataSpace – the Common Data Space for Ireland

A high-level overview of the positioning and approach of IRL-DataSpace (the Common Data Space for Ireland) is illustrated in .

- IRL-DataSpace would serve as a national-level single point of entry for all R&I-focused users, projects and activities that want to **Find** and **Access** R&I-relevant data (open and closed).
- IRL-DataSpace would federate Ireland's R&I-related data systems at technical and governance levels in compliance with national, EU and institution-/activity-specific requirements and policies.
  - This will be achieved through the four functional components: Federated Catalogue of Resources; Data & Processing Services; Identity & Trust Management; FAIR Data & Governance. These are discussed later in this section.
  - IRL-DataSpace and the mature data systems will be **Interoperable** through mutually compatible and agreed access gateways based on standards.
  - Datasets in IRL-DataSpace (from the contributing mature data systems) will be **Interoperable** and **Reusable** based on mutually compatible and agreed Data Sharing Standards and Governance Policies as well as their implementations.
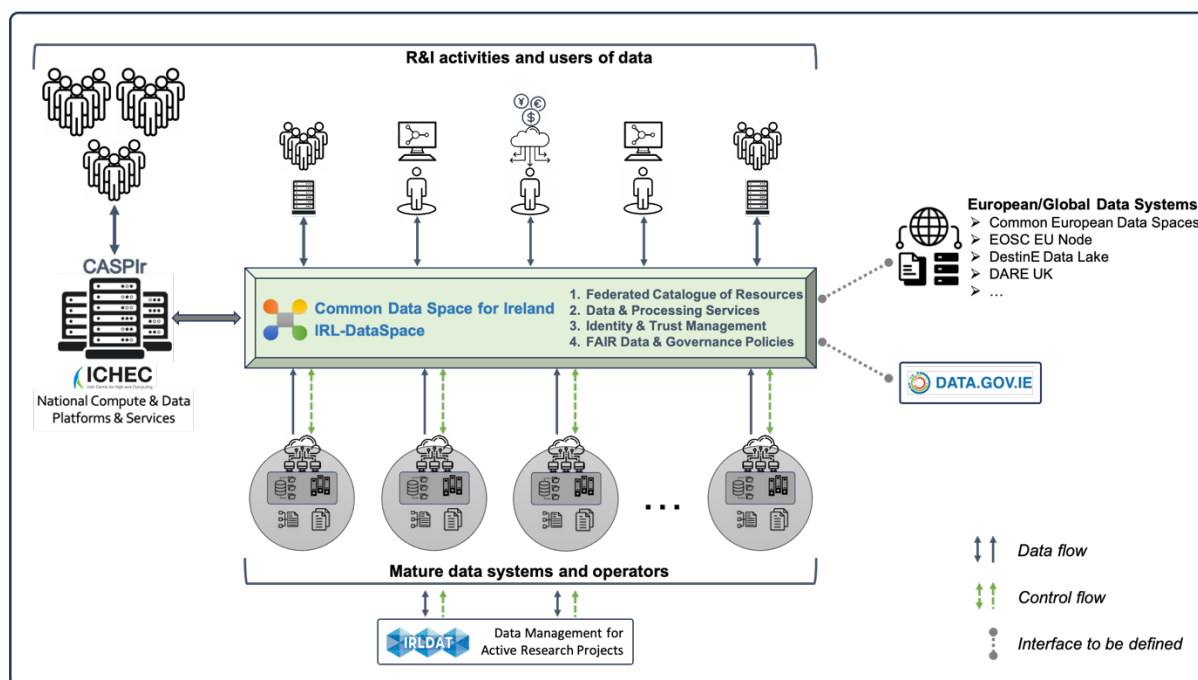
Figure 6: Positioning and approach of the Common Data Space for Ireland (IRL-DataSpace)

- IRL-DataSpace and Ireland's Open Data Portal (data.gov.ie) serve different purposes.
  - The Open Data Portal allows public sector bodies to publish their open data thereby promoting innovation and transparency. This allows Findability and Accessibility.
  - The IRL-DataSpace is to federate a variety of data systems by building and ensuring strict quality assurance for Findability, Accessibility, Interoperability and Reusability of all R&I-related datasets that may be shared by academic, public sector and industry.

Consequently, a suitable interface mechanism should be defined and implemented to ensure that IRL-DataSpace and Ireland's Open Data Portal are **Interoperable**.

- IRLDAT, the pilot platform to manage Active data in research projects in Ireland, will support the mature data systems, which interface with IRL-DataSpace, in managing their active data. IRLDAT is led by HEAnet and implemented collaboratively with ICHEC, UCC, UCD, SETU and EUDAT.
- CASPIr (Compute, Analysis, Simulation Platform for Ireland), the upcoming national Tier-1 data-centric high-performance computing (HPC) system to be provided by ICHEC, will interface with IRL-DataSpace to ensure they are **Findable**, **Acccessible** and **Interoperable** through standardised access gateways and policies for data sharing and governance.
  - This will allow the existing userbase served by the ICHEC National HPC and Data Services across Ireland to seamlessly, efficiently, and securely connect to and use IRL-DataSpace resources from within the CASPIr environment.
  - Other users may also choose to connect directly to IRL-DataSpace from their institutional, personal or public cloud computing environments.

3.2. IRL-DataSpace – its functional elements

As outlined in Section 3.1, IRL-DataSpace would federate with Ireland's and relevant European R&I-related data systems operated by a variety of stakeholders at national, institutional, and discipline levels. This will be achieved through the implementation of the four functional components illustrated in Figure 7.

The four functional components are a combination are technical tools and processes that are implemented and used by those contributing mature data systems (data system operators, data holders, data governors, R&I activities, researchers) in tandem with IRL-DataSpace:
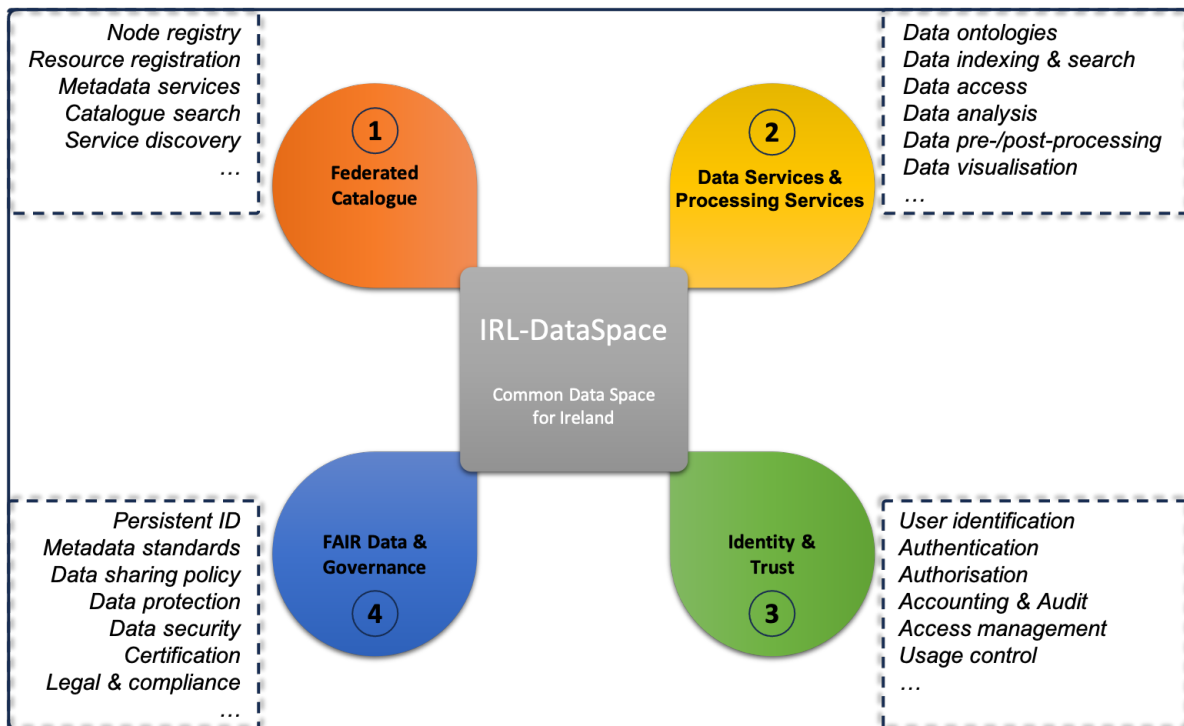
Figure 7: Functional elements of the Common Data Space for Ireland (IRL-DataSpace)

- **FAIR Data & Governance:** In order to synergise the maturation of all data systems that federate via IRL-DataSpace, a common baseline of FAIR-ness principles and governance policies have to be agreed upon and implemented in an interoperable way. This would include persistent identification of data, standardised core metadata, policies and processes for intra- and inter-institutional data sharing, data protection, security, certification (of systems and the data it offers in the federation), and legal and compliance terms for participation in the federation.

- **Federated Catalogue of Resources:** This enables matching between providers and consumers through self-descriptions of resources and services that are offered by the mature data systems via IRL-DataSpace in a standardised and transparent way by listing them in a searchable catalogue each with unique identifiers. Therefore, it facilitates the federation and reuse of these components. Self-descriptions will be protected by digital signatures, which will ensure verification of credentials that are contained within it. The catalogue will be nested and cascaded to ensure uniqueness of identities and resources across the different systems in the federation.

- **Data Services & Processing Services:** Software implementation of interfaces, libraries and tools for standardised, transparent and optimised search, indexing, access, ananlysis, processing and (optional) visualisation of the data shared via IRL-DataSpace. While a baseline of generic interfaces/libraries/tools should be implemented for all data, there exist customisations to address performance efficiency and domain-specific aspects in certain disciplines (such as Earth Systems, Geographic Information Systems, etc.).

- **Identify & Trust Management:** Unique identification of personnel is essential for authentication, authorisation, accounting and auditing of access/use, implement access control and usage policies. IRL-DataSpace will use existing AAI (Authentication and Authorisation Infrastructure) with accounting and auditing methods, rather than maintain them directly.

Figure 7 also lists a non-exhaustive list of core functional elements for each component that would be implemented through tools and processes. It is essential to note that the implementation of these functional components and its constituent elements is a collaborative effort between the IRL-DataSpace developers/operators and the mature data system providers.

Based on this, it is essential to note that

- Some of these IRL-DataSpace functionalities are cross-cutting and agnostic of specific institutions or disciplines. These include
  - in FAIR Data & Governance: core elements of metadata standards, data protection regulations, and certification of resources.
  - in Federated Catalogue of Resources: node registry, resource registration, core functionalities of metadata services, catalogue search, and service discovery.
  - in Identity & Trust Management: persistent ID, federated AAAI (authentication, authorisation, accounting), audit.
  - in Data Services & Processing Services: core elements of data search, data access.

  Such cross-cutting functionalities are often mission-critical for the system and its production-level operation. Hence, they will have to be operated and delivered by national-level providers such as HEAnet and ICHEC, in close partnership with members in the federated system of systems. ICHEC and HEAnet have existing resources as a part of their mandated national digital infrastructure and services which can be leveraged and extended to implement a number of these core cross-cutting functional elements.

- Some of these IRL-DataSpace functionalities will be defined by and remain customised to practices and technological solutions at institutional- and/or discipline-levels. These include
  - in FAIR Data & Governance: customised elements of metadata standards, data sharing policies, data security requirements, legal & compliance policies.
  - in Federated Catalogue of Resources: customised functionalities of metadata services.
  - in Identity & Trust Management: access management, usage control.
  - in Data Services & Processing Services: thematic specialisations/optimisations for data indexing, search and access, data pre-/post-processing, data analysis and visualisation.

This level of detail and precision to the functional elements of IRL-DataSpace and its federated system of systems is essential for the Irish R&I-related data ecosystem to be FAIR with national-level mature data systems, Ireland's Open Data Portal, regional and international data spaces and systems across a wide range of disciplines (including but not limited to the Common European Data Spaces, Copernicus Data Services, Destination Earth Data Lake & Core Platform, European Weather Cloud, Integrated Carbon Observation System and other R&I-relevant data systems).

3.3. IRL-DSSC – the Data Space Support Centre for Ireland

It is proposed that a national-level technical body – a Data Space Support Centre for Ireland (IRL-DSSC) – be established to bridge, support, provide technical expertise and synergise the four stakeholder categerties: data system operators & governors, R&I activities, FAIR data frameworks & programmes, and policy-makers & government stakeholders. It is proposed that the IRL-DSSC be co-led by HEAnet and ICHEC as national providers of digital infrastructures and services across networking, compute and data management technologies.

The IRL-DSSC would fill an existing gap in the Irish ecosystem by providing trusted neutral national-level technical coordination and support, and bridging national and international actors and beneficiaries for effective implementation, operation and sustenance of IRL-DataSpace. Figure 8 illustrates the positioning of IRL-DSSC, including prospective working groups that may be constituted for specific functional elements of IRL-DataSpace.
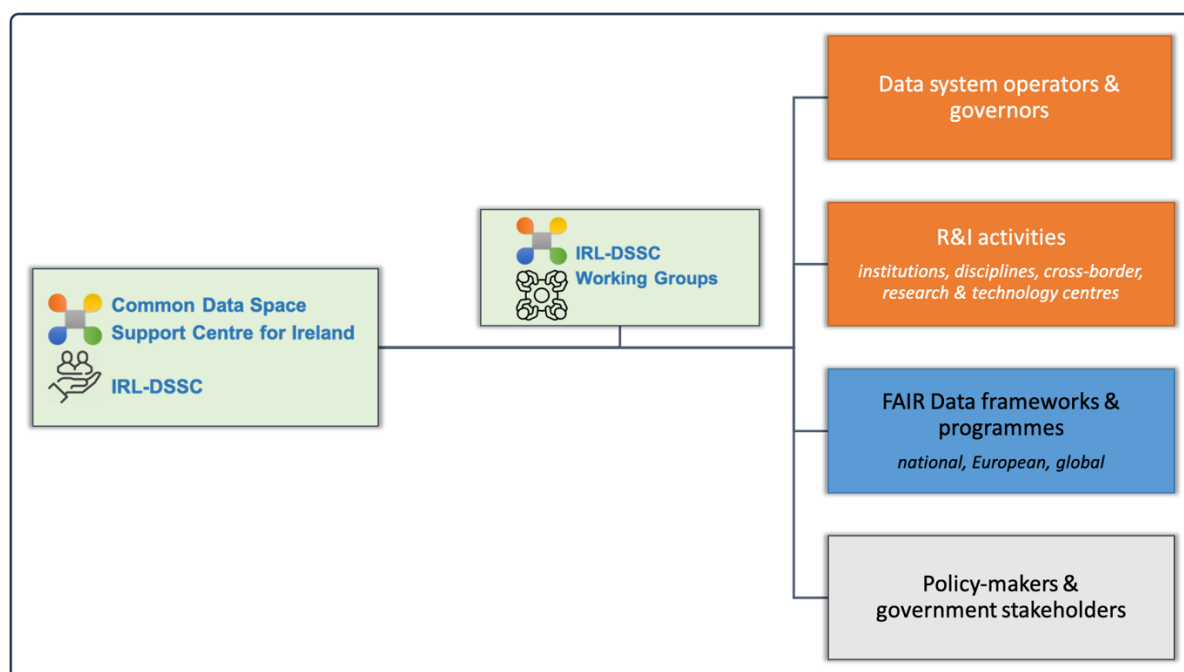
Figure 8: Common Data Space Support Centre for Ireland (IRL-DSSC)

The IRL-DSSC and its working groups will bridge and support the four categories of stakeholders in order to define the specifications, develop, operate and continually expand IRL-DataSpace:

- **Data system operators & governors:** Data systems across Ireland that are relevant and essential for R&I activities at institutional, domain, national and regional levels need to collectively increase, accelerate and align their maturity and FAIR-ness.

  This requires mentorship, agreement, adoption and implementation of consistent and standardised policies, processes, practices and software interfaces/tools to ensure interoperability across all their technical and personnel components.

  IRL-DSSC will serve to remove current siloes between R&I-related data systems, institutions and activities through well-defined interoperable data sharing policies, practices and tools.

- **R&I activities:** Data owners and providers need to improve the quality and standards of their data management for consistent FAIR-ness of the data that they choose to share (open and closed). This requires both individual researchers as well as the broader R&I activities – institutions (e.g. RPOs, state agencies), research centres (e.g. under SFI), technology centres (e.g. under EI), disciplines, projects – to have access to technical and practical expertise for consistent implementation and practice of FAIR data management and sharing.

  This will ensure that R&I activites share their data adhering to FAIR, transparent, proportionate, non-discriminatory, secure and privacy-preserving requirements using well-defined and trustworthy technical and data governance mechanisms.

- **FAIR Data frameworks & programmes:** There are a number of fora, programmes and activities at national, European and global levels that define and support the implementation of FAIR Data management systems and communities.

  These include the National Open Research Forum (NORF), Common European Data Spaces, International Data Spaces, EU Destination Earth Initiative, European Open Science Cloud (EOSC), ESA Copernicus Services, European Weather Cloud, Data and Analytics Research Environments UK (DARE UK), GAIA-X, Big Data Value Association (BDVA), European Technology Platform for HPC (ETP4HPC), and equivalent initiatives in other EU States.

  It is essential for IRL-DataSpace to leverage and contribute to the guidelines, templates and implementations defined in these activites that are most relevant for Ireland.

- **Policy-makers & government stakeholders:** It is essential for IRL-DSSC and IRL-DataSpace to take prioritised national-level impact objectives from policy-makers in government departments and

mission agencies as they are both the (co-)funders of R&I activities, participants and key beneficiaries of a synergistic data ecosystem and the data-driven R&I outcomes.

IRL-DSSC should inform policy-makers and funders of the support required to develop and sustain the federated data ecosystem, and demonstrate the outcomes and their impact.

Most importantly, all participant data systems, their R&I activities and respective systems administration and data governance teams need support and mentorship to collaboratively synergise and mature their data systems to be a part of IRL-DataSpace. While the minimum requirements for this are yet to be formally defined, they will include adoption, definition and adherence to standards, definition of appropriate Service Level Agreeements (SLAs) to assure a minimum level of quality of service, building the national-level nested and cascading catalogue of resources (also categorisable based on mission-critical, best-effort or experimental resources). The IRL-DSSC will engage with academia, enterprises, public sector to ensure stronger data governance for security, sensitivity and privacy in the access and use of open/closed FAIR Data for collaborative R&I within and across disciplines.

Build a Common Data Space for Ireland (IRL-DataSpace) with a Support Centre (IRL-DSSC)

- IRL-DataSpace will serve as a national-level single point of entry to R&I-relevant FAIR data (open and closed) for all R&I-focused users, projects and activities.
- IRL-DataSpace will federate Ireland's R&I-related data systems, owners and providers at technical and governance levels in compliance with national, EU, institution-/activity-specific requirements and policies.
- IRL-DSSC will fill the existing gap by providing trusted neutral national-level technical coordination and support to bridge data system operators, data governors, R&I activities, relevant FAIR data programmes, funders and policy-makers.
- IRL-DataSpace and IRL-DSSC together will collectively increase, accelerate and align the maturity and FAIR-ness of data systems and datasets across Ireland that are relevant and essential for R&I activities at institutional, domain, national and regional levels.
- It is proposed that the IRL-DSSC be co-led by ICHEC and HEAnet as national providers of digital infrastructures and services across networking, compute and data technologies.

---

4. Relevant frameworks and activities for data management

The efforts by IRL-DSSC to build, operate and support IRL-DataSpace needs to leverage and collaborate with relevant existing activities at national and global levels. These include, but are not limited to, the following categories:

Frameworks for implementing and/or federating data management systems, which may be agnostic or specific to different disciplines or regions. The most relevant that are currently being defined or developed are listed in Table 1.

R&I-related activities in Ireland, which span prominent disciplines that have emphasised the data-centric nature of their work and ecosystem. Some of these are developing and/or refining their data management systems towards better FAIR-ness and federation. A few exemplars include (bot are not limited to) climate sciences, environmental sciences, agriculture & food, geosciences, life sciences, genomics, urban sciences, renewable energy, physics and material sciences.

Relevant frameworks	Description
Common European Data Spaces	Driven by the European Commission, this EU funded initiative brings together relevant data infrastructures and governance frameworks in order to facilitate data pooling and sharing. Currently, Common European Data Spaces are being defined and/or developed across 14 sectors/domains. https://digital-strategy.ec.europa.eu/en/policies/data-spaces

Common Data Space for Ireland
IRL-DataSpace

Relevant frameworks	Description
Towards European Health Data Space (TEHDAS)	The TEHDAS Joint Action developed European principles for the secondary use of health data for research and innovation. TEHDAS included partipation from 25 European countries, including Ireland's Department of Health and the Health Research Board (HRB). https://tehdas.eu/
European Open Science Cloud (EOSC)	EOSC champions FAIR research data management and application to ensure scientists have access to data-driven science. For this, EOSC aims to develop a "Web of FAIR Data and Services" for science in Europe. In 2022, the Horizon Europe issued a call to develop the first EOSC node (EOSC EU node) to bootstrap the development of the federation of EOSC nodes. https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud
EU Destination Earth Initiative (DestinE)	DestinE is a flagship initiative of the European Commission to develop a highly-accurate digital model of the Earth (a digital twin of the Earth) to model, monitor and simulate natural phenomena, hazards and the related human activities. DestinE will allow users to access thematic information, services, models, scenarios, simulations, forecasts and visualisations. Key components are DestinE Core Platform (HPC & cloud computing), Data Lake (including data federation), Digital Twin Engine and Digital Twin models. https://destination-earth.eu/
GAIA-X	GAIA-X is an initiative that develops, based on European values, a digital governance that can be applied to any existing cloud/ edge technology stack to obtain transparency, controllability, portability and interoperability across data and services.While GAIA-X has strong enterprise participation, it is not a market operator, nor does it aim to operate directly or exclusively any of the services required by the governance. https://gaia-x.eu/
Earth Systems Grid Federation (ESGF)	ESGF is a Peer-to-Peer (P2P) collaborative enterprise system that develops, deploys and maintains software infrastructure for the management, dissemination, and analysis of model output and observational data. ESGF's primary goal is to facilitate advancements in Earth System Science. It is an interagency and international effort. Ireland's CMIP data contributions to the UN IPCC reports by Met Éireann and EPA are hosted and operated in an ESGF node by ICHEC. https://esgf.llnl.gov/, https://wcrp-cmip.org/nodes/esgf-ichec-ie/
Copernicus Data Services	Copernicus is the Earth Observation (EO) component of the EU Space programme, looking at our planet and its environment to benefit all European citizens. It offers information services using satellite EO and in-situ (non-space) data for atmosphere, marine, land, climate change, security and emergency. A federation of DIAS (Data and Information Access Services) platforms provide centralised access to Copernicus data and processing tools. https://www.copernicus.eu/en/access-data/dias
European Weather Cloud	The European Weather Cloud is driven by ECMWF and EUMETSAT as a distributed platform to develop and operate European meteorological applications, by providing and consuming weather-related data and services. https://www.europeanweather.cloud/
Earth Observation Data Cube	Earth Observations (EO) data have exceeded the petabyte-scale and are increasingly freely and openly available from different data holdings. EO Data Cubes (EODC) are a new paradigm revolutionizing the way users can interact with EO data by facilitating access, processing and use of large spatio-temporal EO data in extremely efficient and optimised analysis ready formats. Different EODC implementations are currently operational such as Euro Data Cube, Open Data Cube, Digital Earth Australia, among others. https://eo4society.esa.int/resources/euro-data-cube/

Relevant frameworks	Description
Data Access, Storage, Sharing, Linkage (DASSL), Ireland	DASSL developed and tested a proof-of-concent technical e-infrastructure to demonstrate supporting health data access, storage, sharing and linking for secondary usage (research and innovation purposes). It's purpose is to inform the roll-out of a national service based on the proposed DASSL model. DASSL was funded by the HRB, led and implemented by ICHEC in partnership with the RCSI, HSE and St. James's Hospital (TCD). https://www.hrb.ie/funding/approved-awards/dassl/
IRLDAT, Ireland	IRLDAT is a pilot for a national shared storage service for Active data in research projects. It is expected that a production service will be offered in 2025. It will be a centrally managed services available to all RPOs in Ireland. It will be based on a subset of services from EUDAT for the active phase of research projects. IRLDAT is led by HEAnet and and implemented collaboratively with ICHEC, UCC, UCD, SETU and EUDAT. https://b2drop.irldat.ichec.ie/apps/dashboard/
Australian Research Data Commons (ARDC)	ARDC enables Australian researchers and industry access to national e-Infrastructure, skills, and data collections. ARDC partners with the research community and industry to build leading-edge digital research infrastructure to provide Australian researchers with competitive advantage through data. https://ardc.edu.au/
National Data Infrastructures (NDI), Czechia	The NDI in Czechia is a common platform for sharing, managing, accessing data and computing resources for research purposes. NDI will support both scientific and multidisciplinary research activities, and will cover a wide range of scientific fields and disciplines. Implementing NDI includes interconnecting existing thematic/disciplinary data repositories into a common infrastructure, and upskilling academic and research institutions that are introducing research data management. https://www.eosc.cz/en
National Research Data Infrastructure (NFDI), Germany	The NFDI aims to systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally. The NFDI brings multiple stakeholders together in a coordinated network of consortia to provide science-driven data services to research communities. https://www.dfg.de/en/research-funding/funding-initiative/nfdi
CSC Research Data Management (RDM) Competence Centre, Finland	The RDM at CSC in Finland provides data management service to research groups during and after the active phase of a research project. The CSC RDM Competence Centre promotes and supports skills and competence development of research data management. The aim is to foster Open Science and FAIR data management with adequate knowledge and best practices for managing, reusing, sharing and analysing research data. https://research.csc.fi/data-management
Data and Analytics Research Environments UK (DARE UK)	DARE UK was established by UK Research and Innovation (UKRI) to design and deliver a coordinated and trustworthy national data research infrastructure to support cross-domain research for public good. It covers all types of sensitive data relating to people, including data about education, health, the environment, welfare and much more. DARE UK aims to establish trusted research environments (TREs) for fast, safe and efficient sharing, linkage and advanced analysis of data, where it is legal and ethical to do so. It will enable UK researchers and innovators to securely and efficiently use linked datasets, modern digital platforms, tools, techniques and skills. https://dareuk.org.uk/
EMPOWER: Data Governance, Ireland	EMPOWER develop systems and methods to enable and support innovative data governance in the real-world application of data-driven solutions. https://empower-datagovernance.ie/

Table 1: Frameworks for implementing and/or federating data managment systems

5.   An All-Island Perspective

There are compelling reasons why IRL-Dataspace supported by IRL-DSSC should include provision for bidirectional sharing of data, platforms and services with Northern Ireland. These include areas in which all-island cooperation is essential such environmental research, and areas in which shared services are already operational. Research collaboration between the Republic and North is increasing rapidly, with Northern Ireland's position at the interface between the United Kingdom and the European Union providing unique opportunities. Working across the border is potentially challenging because of the separate funding mechanisms that are available on either side of the border, and also by potential future regulatory divergence on data protection and other related issues. These are ameliorated by a willingness to collaborate that has seen research funding streams emerge that specifically promote and enable all-island research.

The following are a few major cases of ongoing all-island R&I activities:

- The Climate + Biodiversity and Water Co-Centre is funded by Science Foundation Ireland, Northern Ireland Department of Agriculture, and the Environment & Rural Affairs and UK Research & Innovation, with industry partners[14]. The co-centre brings together over 60 leading researchers from 14 academic partner institutions in Ireland, Northern Ireland and Great Britain. These researchers will work to deliver the transformative change urgently needed to tackle the climate, biodiversity and water crises impacting the two islands – and the wider world. ICHEC and NI-HPC participate in the co-centre to define the data management requirements and specifications particular to the research carried out. A pilot Climate+Biodiversity and Water Data Space is planned to be developed.
- The North-South Research Programme is run by HEA on behalf of DFHERIS and provides more than €50 million of funding for over 60 collaborative projects spanning a diverse range of areas of mutual interest such as food fraud, recycling of agricultural waste, peacebuilding, and a range of medical research projects.[15]
- The US-Ireland Research and Development Partnership promotes and supports collaborative research and development between the US, Ireland, and Northern Ireland, with funding drawn from research funders in each of the three jurisdictions[16]. Research supported by this scheme is within the seven thematic priority areas and either be transferrable to the market, or lead to improvements in healthcare and disease prevention. Supported projects are in areas that include genomics, nutrition, diabetic retinopathy, and bacterial infection.

An exemplar potential use-case for all-island data services is the proposed All-Island Cancer Institute[17]. With up to 50% of residents on the island developing cancer at some point during their lifetime, there is a need to bring together the island's research expertise and facilities to work together and to translate lab science into clinical care. Effective collaboration will require seamless sharing and integration of data, with federated access to shared systems enabling data to be accessed when and where it is needed.

6.   Recommendations to build IRL-DataSpace and IRL-DSSC

IRL-DataSpace and IRL-DSSC will deliver R&I impact across several disciplines in environment, soceity, economy at national and international levels, by dealing with grand challenges and informing better public policy-making. This is consistent with the Programme for Government – *Our Shared Future* – as highlighted in the Impact 2030 Strategy. IRL-DataSpace and IRL-DSSC are an essential component for Ireland to deliver these key national and European strategies.

---

[14] https://www.sfi.ie/research-news/news/two-new-research-centres/

[15] https://hea.ie/policy/research-policy/north-south-research-programme/

[16] https://research.hscni.net/us-ireland-rd-partnership-awards-portfolio

[17] https://research.ie/what-we-do/loveirishresearch/blog/project-spotlight-all-island-cancer-research-institute/

To R&I activities conducted by institutions, research centres, projects or researchers, we highlight that it is essential to coordinate, robustly manage and govern datasets to mature them to open and FAIR Data, facilitated by the IRL-DSSC (Common Data Space Support Centre for Ireland).

To data systems operators, data holders and governors, we call out for their participation to mature the FAIR-ness of their systems and data in a national-level federation based on standardised and FAIR technologies and practices. The IRL-Data Space (Common Data Space for Ireland) is the solution for federation coordinated by the IRL-DSSC (Common Data Space Support Centre for Ireland).

To policy-makers and funders, we emphasise that
- Ireland's ecosystem of diverse data systems will need to be unified as a federated system of systems based on FAIR principles, practices and technologies. This requires incentivisation for the participant R&I activities, data systems operators and data holders to engage.
- Building the IRL-DataSpace (Common Data Space for Ireland) will federate Ireland's R&I-related data systems, owners and providers at technical and governance levels in compliance with national, EU, institution-/activity-specific requirements and policies.
- Establishing the IRL-DSSC (Common Data Space Support Centre for Ireland) would lead the development of IRL-DataSpace through collaboration with the data system operators, data holders, governors and R&I activities. The IRL-DSSC will an fill the existing gap by providing trusted neutral national-level technical coordination and support to bridge data system operators, data governors, R&I acvities, relevant FAIR data programmes, funders and policy-makers.
- This paper on IRL-DataSpace complements the "Vision to Implement the European Open Science Cloud (EOSC) in Ireland" is outlined by HEAnet in consultation with ICHEC and the consortium of Irish Research Libraries (IReL). These two papers are aligned and linked to enable the connection of Data Spaces to EOSC.

It is proposed that IRL-DSSC be co-led by ICHEC and HEAnet as the national providers of digital infrastructures and services across computing data and networking infrastructure and technologies, with sponsorship from relevant government departments and participation of relevant stakeholders. For this, ICHEC and HEAnet have available the expertise that can be leveraged. Futhermore, ICHEC and HEAnet have existing resources as a part of their mandated national digital infrastructure and services which can be leveraged and extended to implement a number of these core cross-cutting functional elements.

Figure 9 illustrates the proposed roadmap (2024-2027) for building the IRL-DataSpace supported by the IRL-DSSC, its working groups and all relevant stakeholders. The roadmap includes
- briefings and consultations to onboard all stakeholders,
- focused white papers to establish technical specifications for IRL-DataSpace, requirements for different R&I disciplines, and all-island joint activities,
- the finalisation and support required from policy-makers and funders, and
- development and operationalising IRL-DataSpace as a National Service by 2027.

Common Data Space for Ireland
IRL-DataSpace

Step 1: Scoping & Positioning

Sep'23

May'24

ICHEC-HEAnet mini-workshop on data ecosystem in Ireland

IRL-DataSpace Position Paper Published on Zenodo Submitted to DFHERIS

Step 2: Briefings & Consultations

Briefings to IE Government Departments & Agencies

Consultation workshop with European experts

Establish IRL-DSSC, Working Groups, Terms of Reference

Aug'24 – Oct'24

Aug'24 – Dec'24

IE workshop with data system operators, providers, governors

IE workshop with R&I activities & researchers

Step 3: Design & Specification

Jan'25 – May'25

Prepare & publish "IRL-DataSpace Technical Specification"

Prepare & publish "IRL-DataSpace for R&I Disciplines"

Prepare & publish "IRL-DataSpace for All-Island R&I Activities"

Step 4: Finalisation & Funding

Jun'25

Jul'25 – Oct'25

Finalise IRL-DataSpace Phase-1 development timeline & select Phase-1 participants

IRL-DataSpace Phase-1 development approved & funding secured

Step 5: Development & Operationalisation

Oct'25 – Dec'26

Oct'26 – Dec'26

IRL-DataSpace National Service 2027 onwards

ICHEC, HEAnet & participants develop & demonstrate IRL-DataSpace Phase-1

Review IRL-DataSpace Phase-1 Roadmap & funding secured for IRL-DataSpace National Service
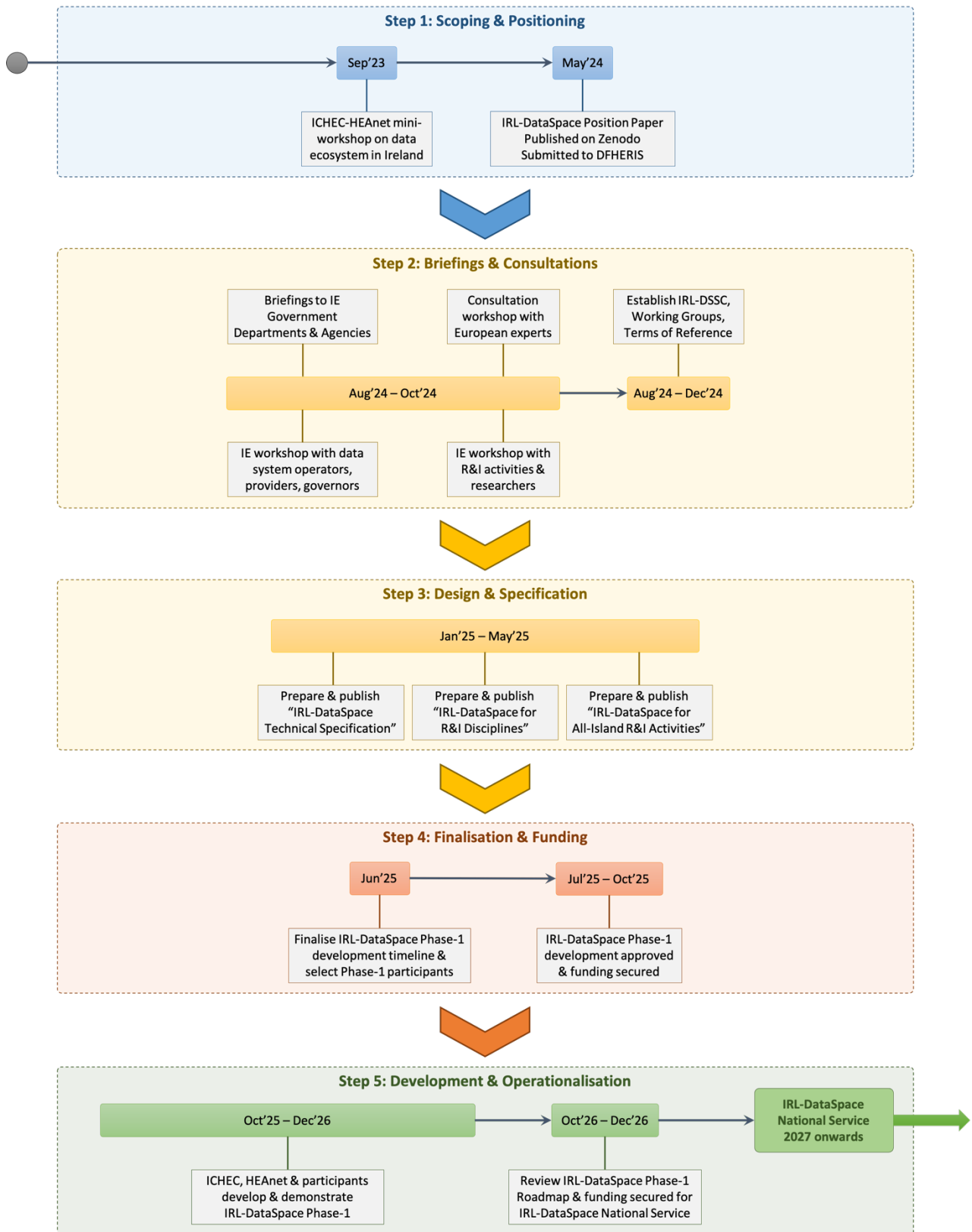
Figure 9: Roadmap and timeline to operationalise IRL-DataSpace and IRL-DSSC