Deliverable 6.2

EHDS Semantic interoperability framework

Recommendations to enhance interoperability within HealthData@EU- a framework for semantic, technical and organisational interoperability

21 December 2022

This project has been co-funded by the European Union's 3rd Health Programme (2014-2020) under Grant Agreement no 101035467

0 Document info

0.1 Authors

Author	Partner
Francisco Estupiñán-Romero	Institute for Health Sciences in Aragon, Spain
Enrique Bernal-Delgado	Institute for Health Sciences in Aragon, Spain
Micaela Comendeiro-Mälloe	Institute for Health Sciences in Aragon, Spain
Persephone Doupi	Finnish Institute for Health and Welfare (THL), Finland
Juan Gonzalez-García	Institute for Health Sciences in Aragon, Spain
Mari Mäkinen	Finnish Institute for Health and Welfare (THL), Finland
Natalia Martínez-Lizaga	Institute for Health Sciences in Aragon, Spain
Malou Munkholm	Central Denmark Region, Denmark
Annelise Nyvold Lundbye	Central Denmark Region, Denmark
Carlos Telleria-Orriols	Institute for Health Sciences in Aragon, Spain

0.2 Keywords

Keywords	TEHDAS, Joint Action, Health Data, Health Data Space, Data Space, HP-JA-2020-1

Accepted in Project Steering Group on 13 December 2022. Updated on 31 October 2023.

Disclaimer

Copyright Notice

Contributors

Contributor	Institution
Linda Abboud	Sciensano
Bodi Antal	Semmelweis
Roxana Arzideh	Bfarm
Lorien Brenda	French Data Hub
Pablo Chaves de Luis	Spanish Ministry of Health
Kerstin Engelhardt	The Norwegian Directorate for e-health
Yasmin Fonseca	SPMS
Roch Giorgi	Aix-Marseille Université
Zdenek Gutter	MZCR
Magnus Häll	Statistics Sweden
Tala Haddad	INSERM
Daniel Karlsson	National Board of Health and Welfare
Toth Kornel	Semmelweis
Truls Korsgaard	The Norwegian Directorate for e-health
Ana Rath	INSERM
Gaby Wildenbos	NICTIZ

Contents

1 Executive summary

This report aims at providing guidance on the use of selected well-established interoperability standards, ultimately to enhance interoperability in the context of HealthData@EU. These recommendations are one of the pieces of the TEHDAS Data Quality Framework.

Interoperability standards play a significant role in some of the stages of the data life cycle; in particular, in the data preparation process, in the publication of data collections, and in the connection between nodes in the HealthData@EU. This report provides assessment and recommendations standards that:

1) provide a layer of semantic interoperability upon the different taxonomies and vocabularies used when collecting data for primary purposes;

2) facilitate the cataloguing of data sources and data collections in a way that enables findability in a programmatic querying; and

3) allow a communication with preserved meaning between nodes when health data access bodies grant access to data.

Finally, the recommendations of this report are primarily aimed at those institutions that are expected to prepare data for secondary use (e.g., data holders, health data access bodies).

Methodology

After the selection of international well-established standards with potential in secondary use was decided, three observers assessed the principles of interoperability using the Common Assessment Method for Standards and Specifications (CAMSS), which is the European guide for assessing and selecting standards and specifications for eGovernment projects. CAMSS evaluates 16 criteria as openness, transparency, reusability, technical neutrality and portability; generic use and how they address users' needs and expectations; potential for cooperation; and the principles of interoperability in the European Interoperability Framework (EIF).

After assessment, each criterion and the standard itself achieve a score. These scores provide a notion of how well the standards address the interoperability requirements and, consequently, what the strengths and weaknesses are.

Results

This report evaluates nine standards for discoverability, five standards meant to enable semantic interoperability and five aiming the interoperable communication between nodes

Standards for discoverability

Overall, among the studied standards, DCAT-AP and INSPIRE were considered best equipped for data discoverability, as per CAMMS evaluation.

These standards, developed in the context of OpenData and Public Administration, are however not so well equipped to respond to the needs of the research communities. Conversely, standards developed in the context of research communities, that ranked lower in reusability or in the assessment of effectiveness and efficiency, ranked at the top with regard to user-centric approach. This finding highlights the actual relevance of these standards in data discoverability for specific research communities - i.e., data types as biosamples, gene sequences, clinical research data, medical records, etc.

None of the above standards are part of a national framework or national strategy in the context of health data or have been widely adopted as a standard for discoverability. Importantly, DCAT-AP, INSPIRE and CESSDA have developed some mutual collaboration.

Standards enabling semantic interoperability

All the standards analysed got a similar interoperability overall score (around 80%), ranging between 74% in Orphanet standards and 82% in OMOP-CDM.

The adoption of these standards is quite uneven. While in the case of SNOMED CT and LOINC there is wide experience across Europe, the case of Orphanet standards is more limited, and the case of OMOP-CDM is essentially linked to research projects on specific domains, although ranks the highest in the principles for cooperation among public institutions. CDISC-STDM is not used in any of the countries that provided insight.

Importantly, SNOMED CT is being mapped to Orphanet standards, OMOP-CDM, CDISC-SDTM, and LOINC has joined SNOMED CT. In addition, all the ICD and ICD-O is mapped to SNOMED CT, as well as, is being used in the Human Phenotype Ontology and in the GLobal Alliance for Genomics and Health. Finally, SNOMED CT is ISO-IDMP compliant allowing the extension of EMA case safety reports.

Standards meant interoperable communication

Overall, except for IDMP, the standards of communication scored 90% and over in accordance with the CAMSS assessment demonstrating suitability for interoperable communication within the HealthData@EU.
DICOM and HL7-FIHR both scored 99% while ISO-8000-110 reached 96% reflecting a lower score in principles for cooperation across institutions. IDMP´s low figures reflect issues concerning transparency, reusability, security and privacy and a lack of assessment of effectiveness and efficiency.

As per the adoption, DICOM has been widely adopted in all the countries surveyed in this study, and the same largely applies to HL7-FHIR. This is not the case for ISO 8000-110 adopted in just one institution of the surveyed countries. IDMP is used in the communication between manufacturers and EMA, but it is not mentioned at the country level. Importantly both DICOM and HL7-FHIR have set up collaborations between them and with the aforementioned semantic standards. IDMP has focused on SNOMED CT and HL7-FHIR. The level of cooperation is uneven. Finally, DICOM has established cooperation with the Cancer Genome Atlas program in the USA.

2 Glossary

Concept	Explanation
Data repository	The data repository is a large database infrastructure, several databases, that collect, manage, and store data sets for data analysis, sharing and reporting.
Data discoverability	Discoverability is the degree to which a data set or source can be found in a search, a file, a database, or other information systems. Discoverability is related to data publication, metadata documentation, and harmonisation. It is different from accessibility and usability, other qualities that affect the usefulness of a piece of information. In the data discovery phase, the data user looks for the data needed to perform their work (answer a research question and/or make decisions regarding new or existing policies or regulations). Once the search is performed, he or she decides on the feasibility of carrying on their study according to the data found, possibly with the advice of data experts.
European Interoperability Framework (EIF)	The European interoperability framework is a commonly agreed approach to the delivery of European public services in an interoperable manner. It defines basic interoperability guidelines in the form of common principles, models and recommendations.
European Interoperability Reference Architecture (EIRA)	The European Interoperability Reference Architecture (EIRA) is an architecture content metamodel defining the most salient architectural building blocks (ABBs) needed to build interoperable eGovernment systems. The EIRA provides a common terminology that can be used by people working for public administrations in various architecture and system development tasks.
FAIR principles	Principles that ensure that data are prepared to be reused by third parties (R). This requires certain levels of syntactic and semantic interoperability (I) at variable and value level and, machine-readable publication of the meta-data describing the

	datasets where they are collected (F). Accessibility (A) will depend on the level of sensitivity of the data and the specific procedures of the data permit institution.
Governance	Governance is the exercise of political, economic and administrative authority necessary to manage a nation's affairs.
Interoperability	Following the European Interoperability Framework, interoperability refers to a) a full compliance with the legal and ethical provisions in each constituent node, b) an organisation that supports knowledge exchange and software transference across nodes, c) a compatible technological environment that supports the communication between nodes and allows the deployment of the computational tasks, and d) the existence of common data models that enables semantic standardisation across data sources. In a distributed research infrastructure, interoperability is a key feature for its governance and achievements.
National Interoperability Framework (NIF)	A National Interoperability Framework is a commonly agreed approach that cover a set of criteria and recommendations with regard to security, preservation and standardisation to be taken into account by Public Administrations within a country for technological decisions that guarantee a suitable level of legal, organisational, semantic, and technical interoperability of the data, information and services that manage in the exercise of their competences among them and with the citizens.
Openness	The level of openness of a specification/standard is decisive for the reuse of software components implementing that specification. This also applies when such components are used to introduce new European public services.

Organisational interoperability	This refers to the way in which public administrations align their business processes, responsibilities and expectations to achieve commonly agreed and mutually beneficial goals. In practice, organisational interoperability means documenting and integrating or aligning business processes and relevant information exchanged. Organisational interoperability also aims to meet the requirements of the user community by making services available, easily identifiable, accessible and user focused.
Reusability	Reuse means that public administrations confronted with a specific problem seek to benefit from the work of others by looking at what is available, assessing its usefulness or relevance to the problem at hand, and where appropriate, adopting solutions that have proven their value elsewhere. This requires the public administration to be open to sharing its interoperability solutions, concepts, frameworks, specifications, tools, and components with others.
Semantic interoperability	Semantic interoperability ensures that the precise format and meaning of exchanged data and information is preserved and understood throughout exchanges between parties, in other words 'what is sent is what is understood'. In the EIF, semantic interoperability covers both semantic and syntactic aspects: -The semantic aspect refers to the meaning of data elements and the relationship between them. It includes developing vocabularies and schemata to describe data exchanges and ensures that data elements are understood in the same way by all communicating parties; - The syntactic aspect refers to describing the exact format of the information to be exchanged in terms of grammar and format.
Standards facilitating programmatic discoverability and findability	In this work the following standards have been assessed: Beacon, BBMRI-MIABIS, Bio-image archive, CESSDA CMM, DCAT-AP, ECRIN-CRMDR, FAIRSHARING, INSPIRE, PHIRI.
Standards facilitating interoperable communication	In this work the following standards have been assessed: DICOM, HL7 FHIR, IDMP (SPOR), ISO 800-110.

Standards facilitating semantic interoperability	In this work the following standards have been assessed: CDISC SDTM, LOINC, OMOP-CDM, Orphanet standards, SNOMED CT.
Syntactic interoperability	The syntactic aspect refers to describing the exact format of the information to be exchanged in terms of formats, conceptual and logical models, and organization of the information (i.e., variable structure, units, type of data, transformation and validation rules, etc.)
Transparency	Transparency in the EIF context refers to: Enabling visibility inside the administrative environment of a public administration, ensuring availability of interfaces with internal information systems and securing the right to the protection of personal data.
W3C standards	W3C standards define an Open Web Platform for application development that has the unprecedented potential to enable developers to build rich interactive experiences, powered by vast data stores, that are available on any device.

3 Acronyms

Name	**Meaning**
ATC	Anatomical Therapeutic Chemical (ATC) Classification (http://who.int)
BBMRI	Biobanking and BioMolecular resources Research Infrastructure
CAMSS	Common Assessment Method for Standards and Specifications
CDISC SDTM	The Clinical Data Interchange Standards Consortium (CDISC) - Study Data Tabulation Model (SDTM)
CESSDA CMM	Consortium of European Social Science Data Archives
ContSys	International Standard EN ISO 13940:2015

DCAT-AP	Data CATalog vocabulary (DCAT)- Application Profile for data portals in Europe (AP)
DDI	Data Documentation Initiative
DICOM	Digital Imaging and Communications in Medicine
Disco	DDI-RDF Discovery Vocabulary
DQF	Data Quality Framework
DQV	Data Quality Vocabulary
DUO	Data Usage Ontology
ECRIN	European Clinical Research Infrastructure Network
EHDS1	European Health Data Space for primary use (healthcare assistance) - also known as MyHealth@EU
EHDS2	European Health Data Space for secondary use (regulatory, health policy, research and innovation) - also known as HealthData@EU
EIF	European Interoperability Framework
EIRA	European Interoperability Reference Architecture
ELIXIR	The European life-sciences Infrastructure for biological Information
EMA	European Medicines Agency
ERI	European Research Infrastructure
Eurobioimaging	European research infrastructure for biological and biomedical imaging
FAIRSHARING	Findability, Accessibility, Interoperability, and Reusability (FAIR)-Sharing

GSIM	Generic Statistical Information Model
HaDEA	European Health and Digital Executive Agency
HealthData@EU	European Health Data Space for secondary use (EHDS2)
HL7 FHIR	Health Level Seven International (HL7)-Fast Healthcare Interoperability Resources (FHIR)
ICD	International Classification of Diseases
ICF	International Classification of Functioning, Disability and Health (ICF)
ICPC	International Classification of Primary Care
IDMP (SPOR)	Identification of Medicinal Products (IDMP)-Substance, Product, Organisation and Referential (SPOR)
INSPIRE	Infrastructure for Spatial Information in Europe
ISO 8000 110	International Organization for Standardization (ISO)
LOINC	Logical Observation Identifiers Names and Codes
MDR	ISO/IEC 11179 Metadata Registry
MPEG-21 REL	International standard ISO/IEC 21000-5
MyHealth@EU	Electronic cross-border health services in the EU (EHDS1)
NANDA	North American Nursing Diagnosis Association
NIC	Nursing Interventions Classification
NOC	Nursing Outcomes Classification

NPU	Nomenclature for Properties and Units (NPU)
ODRL	Open Digital Rights Language
OMOP-CDM	Observational Medical Outcomes Partnership- Common Data Model
Orphanet ORDO	Orphanet Rare Disease Ontology (ORDO)
PHIRI-HIP	the Population Health Information Research Infrastructure (PHIRI)-Health Information Portal (HIP)
RDF	Resource Description Framework
SKOS	Simple Knowledge Organization System
SNOMED CT	Systematized Nomenclature of Medicine - Clinical Terms
TEHDAS	Towards the European Health Data Space Joint Action
WOL	Web Ontology Language
WP	Work Package
XACML	Extensible Access Control Markup Language

4 Concepts

TEHDAS Data Quality Framework

The TEHDAS data quality framework (DQF) aims at setting up the basis for a trustworthy and reliable secondary use of data and at providing guidance on its implementation.

The TEHDAS DQF builds on the principles of a) trust across data institutions and users; b) transparency in the processing of the data, from the collection to the publication of meta-data; and c) continuous improvement, benchmarking and promotion.

Quality dimensions of major importance in the TEHDAS DQF are:

Dimension	**Definition**
Reliability	How closely it reflects what it was designed to measure and whether this is consistent over time.
Relevance	Meets the needs of users of the EHDS.
Timeliness	Collected within a reasonable period of time and collected/reported on dates agreed, e.g., close to decision makers' time of decision.
Coherence	Consistent over time and across data holders and can be combined and compared with other data sources.
Coverage	The degree to which the data adequately covers the population/event (i.e. representativeness).
Completeness	How complete are the variables?

Although the TEHDAS DQF applies to HealthData@EU (i.e. the European Health Data Space for secondary use), the way data is collected within MyHealth@EU (i.e. the European Health Data Space for primary use) and within the European Research Infrastructures (e.g., ELIXIR, BBMRI-ERIC, ECRIN, Eurobioimaging) is key in the implementation efforts once data is made available for secondary use.

5 Scope of this deliverable

This deliverable is one of the pieces that will compose the TEHDAS Data Quality Framework (DQF) closely interacting with other pieces, in particular, the DQF maturity model.

This deliverable is about providing guidance on the potential use of selected well-established interoperability standards in the context of the HealthData@EU data life cycle. These standards should play some role once data has been made available for secondary purposes.

Standards will play their role at some of the stages in the data-life cycle, as defined in TEHDAS. In particular, in the data preparation process, in the publication of data collections, and in the connection between nodes in the HealthData@EU. More specifically, the recommended standards are expected to: 1) provide a layer of semantic interoperability upon the different taxonomies and vocabularies used when collecting data for primary purposes; 2) facilitate the cataloguing of data sources and data collection in a way that enables findability in a programmatic querying; and 3) allow a communication with preserved meaning between nodes when health data access bodies grant access to data.

Implicit to the selection of standards has been the intention to allow the publication, retrieval and processing of all the potential types of data that will be subject of HealthData@EU - electronic health records, registries, claims data, lab data, image data, data from biosamples, including genes, data from clinical research, etc, and the data models and metadata on those data sources.

The recommendations out of this report will be directed primarily to those institutions that are expected to prepare data for secondary use (i.e., data holders, health data access bodies). Some recommendations may translate into externalities in the health data collection for primary purposes (e.g. an incentive for more detailed recording of clinical data).

Nevertheless, this report is not about providing recommendations on the collection of health data, for example the substitution of taxonomies or vocabularies currently in use for those in this report.

As explained in Milestone 6.2, the conversations along the TEHDAS WP6 participants and the interaction with the EHDS2 pilot working groups recommended extending the scope of this D6.2 to standards for the publication of datasets (meta-data standards), and standards of communication, where syntactic interoperability is key. Finally, DG Santé officers raised a question on the governance of interoperability suggesting the assessment of the standards using the CAMSS tool which implied assessing not just semantic interoperability but all the other layers of interoperability as per the European Interoperability Framework. As the scope of D6.2 was enlarged, the title of D6.2 has also been adjusted to reflect this change.

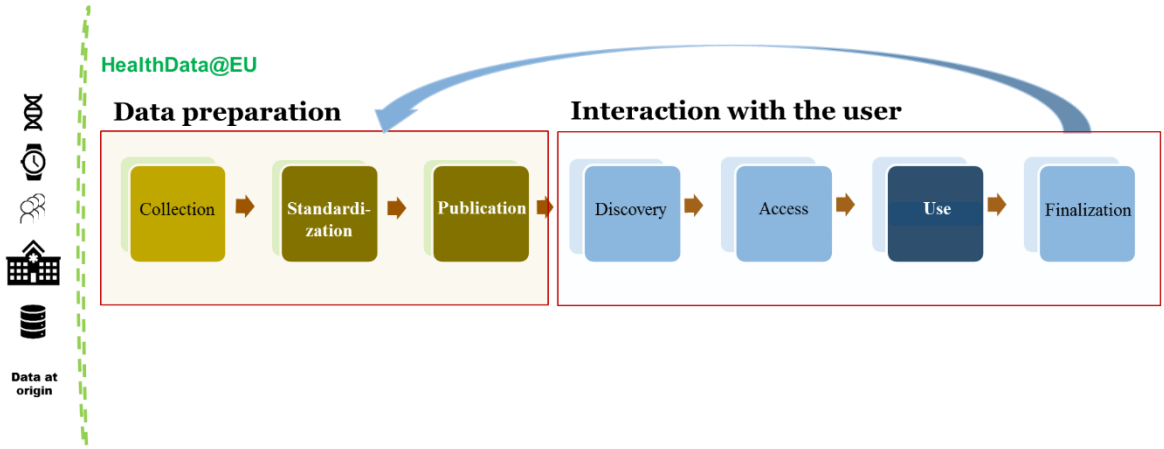
Taking into account the EC requirements for guidance to implement draft EHDS Regulation, D6.3 will describe the building blocks of the TEHDAS DQF - the rationale behind, what is relevant to data quality and utility, who should be in charge at each level (data holders level, HDABs level, SPE level, Users level) and which are the main implementation tasks to govern with which governance tool.

The data-life cycle and the users' journey

The TEHDAS data life cycle and user's journey seek to describe the process that the different actors interacting within the HealthData@EU should follow once data collected for primary purposes is made available for secondary uses.

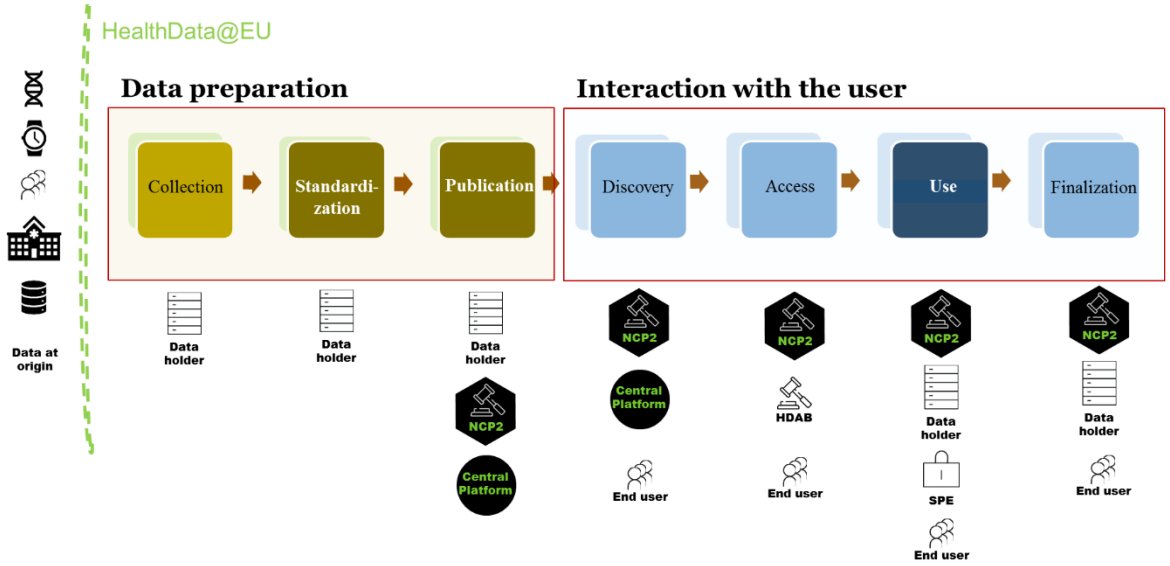
The data life cycle distinguishes between two overarching phases (1) data preparation and (2) interaction with the end user as in figure 1. The former entails the retrieval of data or collection of metadata on primary sources, their preparation for secondary use making them interoperable, and the publication of preparation procedures, data sources and data collections in a way that is easily findable. The latter describes the stages comprising the users' journey, the interaction of the end user with the institutions that may grant access to data; so, once data collections of interest are discovered, how to ask for access permissions, how to access and use the actual data, and how to finalise the use of data including devolution of intermediate outputs and enriched dataset to the data preparation institutions.

Figure 1. Data life cycle



The actors interacting along this journey may be different. In figure 2, relevant actors are depicted according to their roles within the HealthData@EU (This mapping of actors and their role ought to be taken carefully as the legislative proposal regulating the HealthData@EU is under discussion).

Figure 2. Actors within the HealthData@EU data life cycle



The data life cycle has helped identify what services HealthData@EU actors should provide at each stage (see in TEHDAS WP7 deliverable on minimum technical services). In relation to the data quality, figure 3 provides a notion of the different quality elements to be taken into consideration at each stage of the journey.

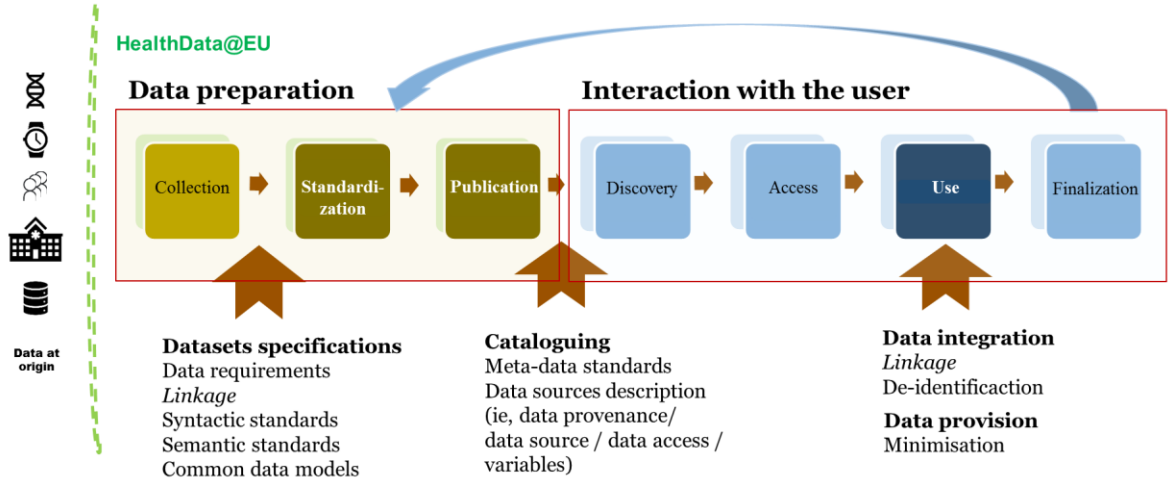
So, in the preparation phase, quality would highly rely on:

1) establishing clear data requirements (*for example, the minimum data to make available, the preparation for data source linkage, the harmonisation of data sources and data collection to be semantically interoperable*);

2) a programmatically interoperable cataloguing of the sources (*implementing meta-data standards that allow a description of the data sources, their provenance, and preparation procedures*).

Once the users have been granted access, the quality will depend on the impact of linkage between data sources, de-identification and minimization procedures before data is made effectively available for use.

Figure 3. Services within the data-life cycle that may have an impact on data quality



The role of interoperability standards in the TEHDAS Data-life Cycle

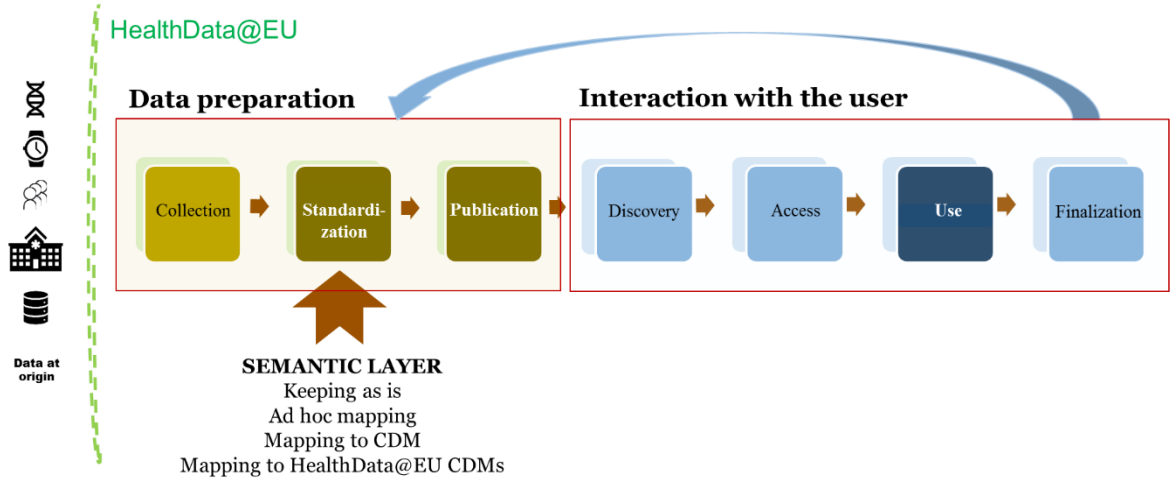
As mentioned earlier, there are a number of stages where standards of interoperability will have a role.

Where to set up a semantic layer

It is expected that data holders (e.g., hospitals, primary care centres, research infrastructures) make their data sources or collections available for secondary use in their original “language”; thus, the controlled vocabularies and taxonomies that they use, for instance, the international classifications of diseases (ICD), the international classification of primary care (ICPC) or the Anatomical, Therapeutic, Chemical classification system (ATC).

Data preparation institutions (i.e., data holders, health data access bodies) are expected to prepare the data to be reusable, thus in a way that is interoperable. In figure 4, there is a notion of what could be the different options to pursue semantic interoperability. Data preparatory bodies a) may decide to keep their data collection as is, using the dictionaries, taxonomies and classification systems in which they code their data sources, b) may decide to prepare *ad hoc* mappings across taxonomies, or c) may decide to use standards meant to build common data models, specific to a number of research queries, or a more general, compatible with the ones recommended by the HealthData@EU.

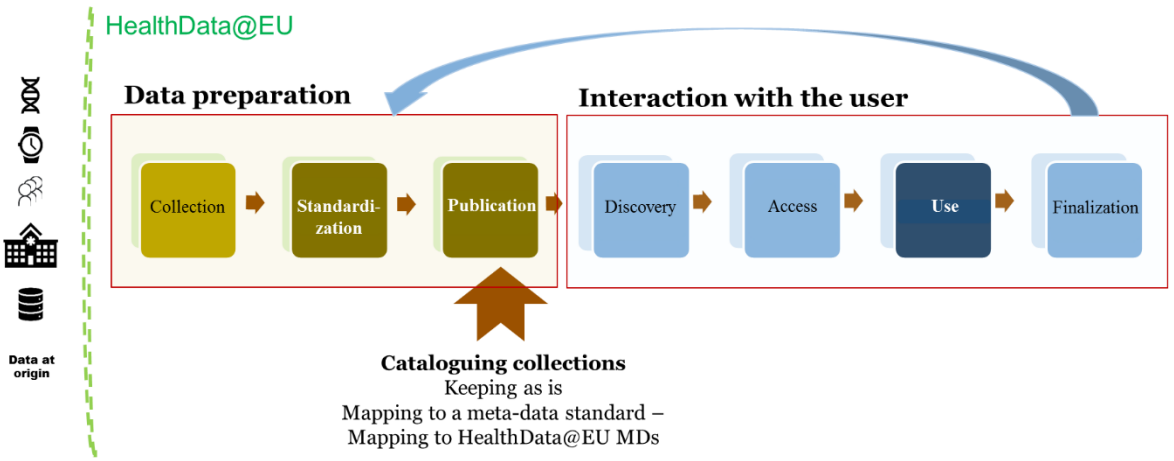
Figure 4. Building a semantic layer within the HealthData@EU data life cycle



Where to set up an Interoperable Catalogue

The institutions in charge of data preparation are expected to publish their procedures and data collections in a way that is programmatically tractable (i.e. both findable and queryable). In this case, what is relevant is the adoption of meta-data standards. In figure 5, there is a notion of where in the data-life cycle and the potential courses of action; thus, data preparation institutions may already have a meta-data catalogue covering part or all of the data sources and data collections that they own (*or are responsible for - stewardship*) and decide to keep it as is; may want to adopt a meta-data standard that allows the interaction with the end users in a programmatic way; or may want to adopt a meta-data standard that would allow programmatically querying across the HealthData@EU (federated querying).

Figure 5. Cataloguing within the HealthData@EU data life cycle

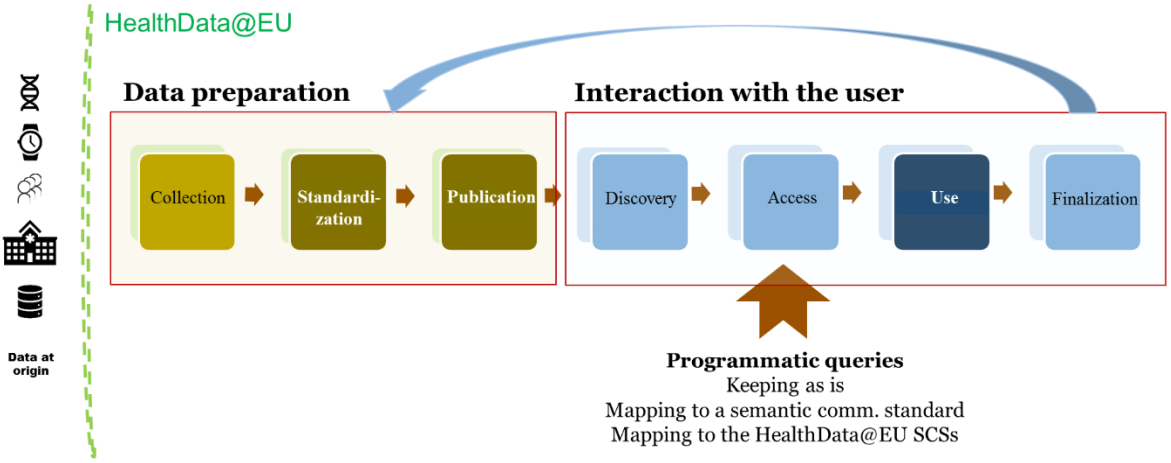


Where to use communication standards

Finally, users' queries for access will be *de facto* sent out to numerous Health Data Access Bodies in a programmatic way. The principle is being able to ask once and get multiple responses. Unlike TEHDAS WP7, where IT colleagues are very much focused on how to technically connect the different actors within the HealthData@EU, this deliverable analyses what communication standards would be of use for a semantic communication across nodes. Figure 6 shows that the application of these standards is expected to happen once access to data is granted.

Health Data Access Bodies a) may think of using their usual communication standards, b) may decide on mapping to an *ad hoc* standard that allows the communication to be semantically interoperable, or c) opt for the recommended semantic communication standards.

Figure 6. Semantic interoperability in the communication within the HealthData@EU



6 Methodology

Identification of relevant standards

The standards collected in this very milestone have been identified via the active participation of WP6 leaders in a number of events and working groups listed in the table below.

Date	Organiser	Content
14/1/21	PHIRI Stakeholders' meeting	Session on the requirement to build federated research infrastructures for a rapid policy response available here

18-19/05/21	TEHDAS Project Forum	Session on data quality and semantic interoperability, agenda available here
18/5/21	EMA	Data Standardisation Strategy stakeholder workshop, presentations available here
11/10/21	EC Workshop	Maximising investments in health research: FAIR data for a coordinated COVID-19 response
29/10/21	TEHDAS Project Forum no 2	Session on Semantic interoperability, data quality assurance
7/12/21	EMA	EU Big Data Stakeholder Forum
2021 1st half	DG Santé	EHDS2 Pilot working groups
12/01/2022	EHealth Network	Reaction to a v.0 list

Once a preliminary list was available, the WP6 coordination team applied the aforementioned roles on the data life cycle and released a second preliminary list for further discussion with WP6 partners. A dedicated slot in a WP6 meeting presenting the approach and the list took place on December 16th, 2021. WP6 partners had the opportunity to provide feedback until December 31st, 2021.

The final list of standards is as follows:

Standards for data discoverability (meta-data standards)

Name	**Domain**	**More information**
Beacon	Genomics	https://beacon-project.io/
BBMRI-MIABIS	Bio-samples	https://github.com/BBMRI-ERIC/miabis
Bio-image archive	Bio-images	https://www.ebi.ac.uk/bioimage-archive/
CESSDA CMM	Social data	https://datacatalogue.cessda.eu/

Name	Domain	More information
DCAT-AP	Public data catalogues	https://ec.europa.eu/isa2/solutions/dcat-application-profile-data-portals-europe_en/
ECRIN-CRMDR	Clinical research	https://ecrin.org/tools/clinical-research-metadata-repository
FAIRSHARING	Digital objects of any kind	https://fairsharing.org/biodbcore/?q=health%20record
INSPIRE	Geo-located data	https://inspire-geoportal.ec.europa.eu/
PHIRI	Population health data	https://www.healthinformationportal.eu/health-information-portal

Standards that enable semantic interoperability

Name	Domain	More information
CDISC SDTM	clinical and pharmaco-epidemiologic studies	https://www.cdisc.org/
LOINC	ontology on lab data	https://loinc.org/ joint-venture with SNOMED CT
OMOP CDM	clinical and pharmaco-epidemiological data model	https://www.ohdsi.org/data-standardization/ https://athena.ohdsi.org/search-terms/terms?domain=Condition&sort=vocabulary_id&order=asc
Orphanet standards	ontology on rare diseases	http://www.orphadata.org/cgi-bin/index.php#ontologies http://www.orphadata.org/cgi-bin/index.php
SNOMED CT	ontology on clinical concepts	https://www.snomed.org/

Standards for interoperable communication

Name	Domain	More information
DICOM	Medical imaging	https://www.dicomstandard.org/

Name	Domain	More information
HL7 FHIR	Electronic medical records	https://www.hl7.org/fhir/ https://wiki.art-decor.org/index.php/Main_Page
IDMP (SPOR)	Medical products	https://www.ema.europa.eu/en/human-regulatory/research-development/data-medicines-iso-idmp-standards/substance-product-organisation-referential-spor-master-data
ISO 8000-110	Data from any master file	https://www.iso.org/standard/78501.html

From the original list, a number of standards were dropped for the following reasons:

Standard	Reason for dropping
ISO23494	In development - no public information released yet
CEDAR	Meta-data tool box on experimental research
EDQM/EMDN	Cataloguing of medical products /medical devices

Methodology of analysis

Once standards were identified, the assessment followed three threads of work: a) interviewing standard promoters; b) evaluating the standards according to CAMMS; and, c) surveying countries participating in WP6. With all the information gathered, a working group within WP6 started the discussion in three workshops, where decisions were made on the final scope, as well as recommendations out of this process.

Interviews with standards' owners

A 2-hour semi-structured interview with the owners or promoters of each Standard in the annexes.

Each interview included the following topics:

- **Background-profile information** of the Standard;
- **Current use** of the Standard;
- **Suitability of the Standard for secondary use** of health data in the context of Healthdata@EU (i.e. policy making, regulation and research purposes);
- **Governance** issues;
- **Implementation barriers, challenges and facilitators**;
- **Requirements for the adoption** – potential supports in the process;
- **Maintenance costs and sustainability**.

CAMSS – as a standard tool

Why CAMSS

The Common Assessment Method for Standards and Specifications (CAMSS) is the European guide for assessing and selecting standards and specifications for an eGovernment project. In this case, the project is the development of the European Health Data Space for secondary use, where setting up the basis for interoperability is key. One of the core elements of interoperability within the EHDS2 has to do with seeking ways for a programmatic discovery of data institutions and data sources of interest, looking for a semantically interoperable way to mobilise the existing data, and allowing an interoperable communication between the actors in the EHDS2.

Within CAMSS, this report has assumed the assessment tools that consider the ***European Interoperability Framework (EIF) perspective.***

Why the EIF perspective

The European Interoperability Framework (EIF) provides recommendations on how to set up interoperable digital public services. The assumption is that the development of the EHDS2 entails setting up interoperable digital public services for the purpose of research and innovation, regulation and policy making across Europe.

CAMSS (EIF) criteria is defined at Toolkit v.5.0.0

Domain	**Criteria**
Core principles	Openness
	Transparency

	Reusability
	Technological neutrality and portability
Generic use and expectations	User centricity
	Inclusion and Accessibility
	Security and Privacy
	Multilingualism
Principles for Cooperation	Administrative simplification
	Preservation of Information
	Assessment of effectiveness and efficiency
Interoperability	Interoperable governance
	Legal interoperability
	Organisational interoperability
	Semantic interoperability
	Technological interoperability

Those criteria were assessed for each standard using the information retrieved from the standard promoters, the analysis of the documentation provided during the interviews and the information gathered from public documents. At least two observers discussed how each standard fulfilled the criteria and contrasted any discrepancy in the interpretation of the criteria with a third observer. Following the CAMSS methodology, each criterion was scored.

Criteria regarding a) subsidiarity of the countries to implement or apply the standards, b) legal interoperability of the standards; and, c) technological interoperability of the standards were considered non-applicable in this assessment. In the first case, because the assessment should be carried out within each country; nonetheless, we have approached this question in

the country surveys asking whether the standards are part of the National Interoperability Framework. In the case of legal interoperability, the assessment should have been done in relation to the current EU regulation still in discussion in the Council and the Parliament; and, in the case of technological interoperability, CAMSS v.5.0.0 deems this layer as a prerequisite for assessment. For these two criteria, after double-checking that all the standards in this study had been implemented in real life following legal and documented technical specifications, we assumed they were legally interoperable and technologically interoperable within the scope of this work.

Country surveys

The country surveys aimed at gaining insight into the experience of the countries with the standards, and fundamentally at inferring difficulties encountered in implementation, governance and sustainability. Three questions were asked:

1) *Is your country familiar (has any experience) with any of the standards considered?*
2) *Which of these standards are widely adopted in your country? And, more specifically, are any of these standards (or specification of them) included in a National Interoperability Framework (NIF) of your country?*
3) *If any of these standards has been adopted in your country (NIF), which have been the main implementation barriers and challenges for their adoption?*

Working group meetings in WP6

- 1st WP6 Working group S&S (04/10/22): discussion of the interview results. The tables with the synthesis of the outcomes of the interviews on interoperability standards/initiatives were presented and the results obtained were discussed.

- 2nd WP6 Working group S&S (31/10/22): scope and structure of Deliverable 6.2. The scope of Deliverable 6.2 and the way to present the results/conclusions for each standard/initiative analysed were discussed.

- 3rd WP6 Working group S&S (21/11/22): discussion of the final draft of Deliverable 6.2: drafted Deliverable 6.2 recommendations were discussed.

7 Results

Section 1. CAMSS results (standard by standard)

1.1 Standards for data discoverability

The standards in this group are: Beacon, BBMRI MIABIS, Bioimage archive, CESSDA CMM, DCAT-AP, ECRIN CRMDR, FairSharing, INSPIRE and PHIRI-HIP. Average scores for each standard are shown in table 1.

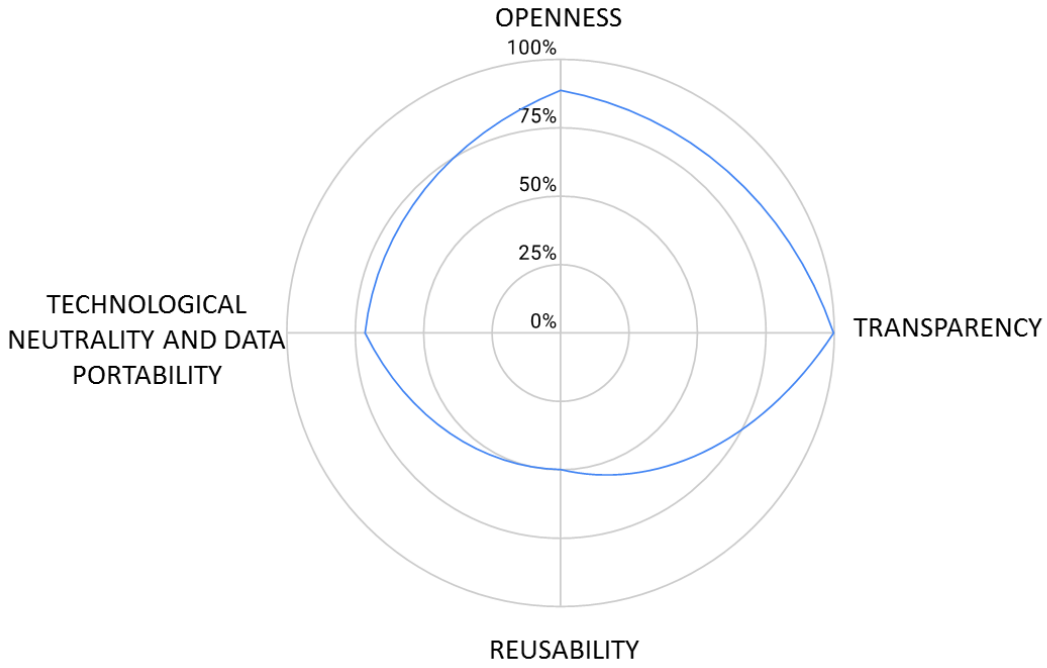
Table 1. Average scores for discoverability standards

CRITERION	BEACON	BBMRI (MIABIS)	BIOIMAGE-ARCH	CESSDA (CMM)	DCAT-AP	ECRIN (CRMDR)	FAIRSHARING	INSPIRE	PHIRI-HIP
CORE INTEROPERABILITY PRINCIPLES	78%	71%	37%	94%	100%	78%	82%	100%	63%
OPENNESS	89%	100%	67%	100%	100%	89%	56%	100%	78%
TRANSPARENCY	100%	100%	25%	75%	100%	75%	100%	100%	25%
REUSABILITY	50%	0%	0%	100%	100%	50%	100%	100%	50%
TECHNOLOGICAL NEUTRALITY AND DATA PORTABILITY	71%	86%	57%	100%	100%	100%	71%	100%	100%
USERS' NEEDS AND EXPECTATIONS	63%	100%	75%	100%	100%	88%	100%	100%	25%
USER-CENTRICITY	100%	100%	100%	100%	100%	100%	100%	100%	100%
INCLUSION AND ACCESSIBILITY	100%	100%	100%	100%	100%	100%	100%	100%	0%
SECURITY AND PRIVACY	50%	100%	0%	100%	100%	50%	100%	100%	0%
MULTILINGUALISM	0%	100%	100%	100%	100%	100%	100%	100%	0%
PRINCIPLES FOR COOPERATION BETWEEN INSTITUTIONS	33%	67%	67%	67%	100%	67%	67%	100%	33%
ADMINISTRATIVE SIMPLIFICATION	100%	100%	100%	100%	100%	100%	100%	100%	100%
PRESERVATION OF INFORMATION	0%	100%	100%	100%	100%	100%	100%	100%	0%
ASSESSMENT OF EFFECTIVENESS AND EFFICIENCY	0%	0%	0%	0%	100%	0%	0%	100%	0%
INTEROPERABILITY LAYERS	71%	96%	88%	96%	100%	88%	79%	100%	38%
INTEROPERABILITY GOVERNANCE	33%	83%	50%	83%	100%	50%	17%	100%	0%
LEGAL INTEROPERABILITY	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ORGANISATIONAL INTEROPERABILITY	50%	100%	100%	100%	100%	100%	100%	100%	0%
SEMANTIC INTEROPERABILITY	100%	100%	100%	100%	100%	100%	100%	100%	50%
TECHNICAL INTEROPERABILITY*	100%	100%	100%	100%	100%	100%	100%	100%	100%
OVERALL SCORE	**61%**	**83%**	**67%**	**89%**	**100%**	**80%**	**82%**	**100%**	**40%**

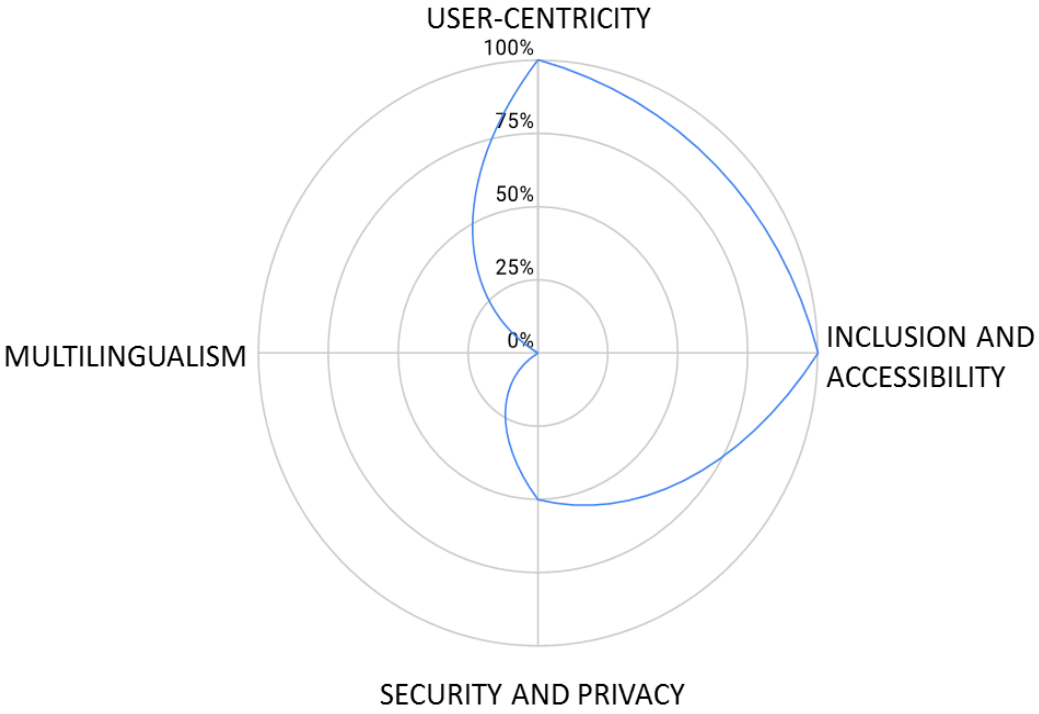
1.1.1 Beacon

1.1.1.1 Results in graphs (in percentages)

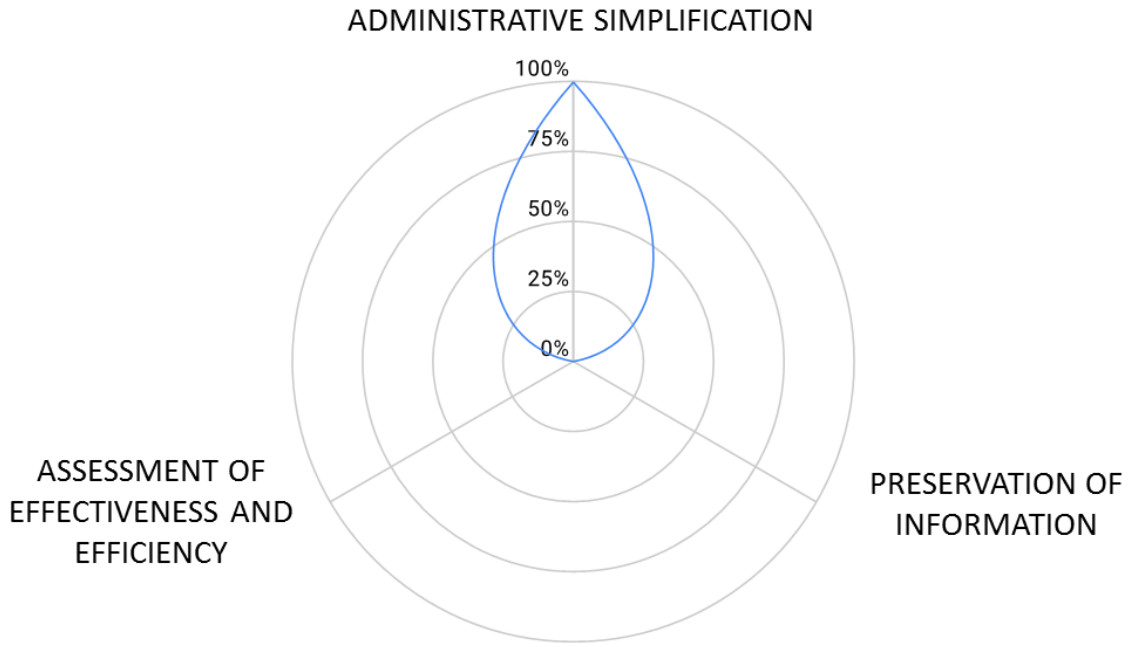
Core interoperability principles



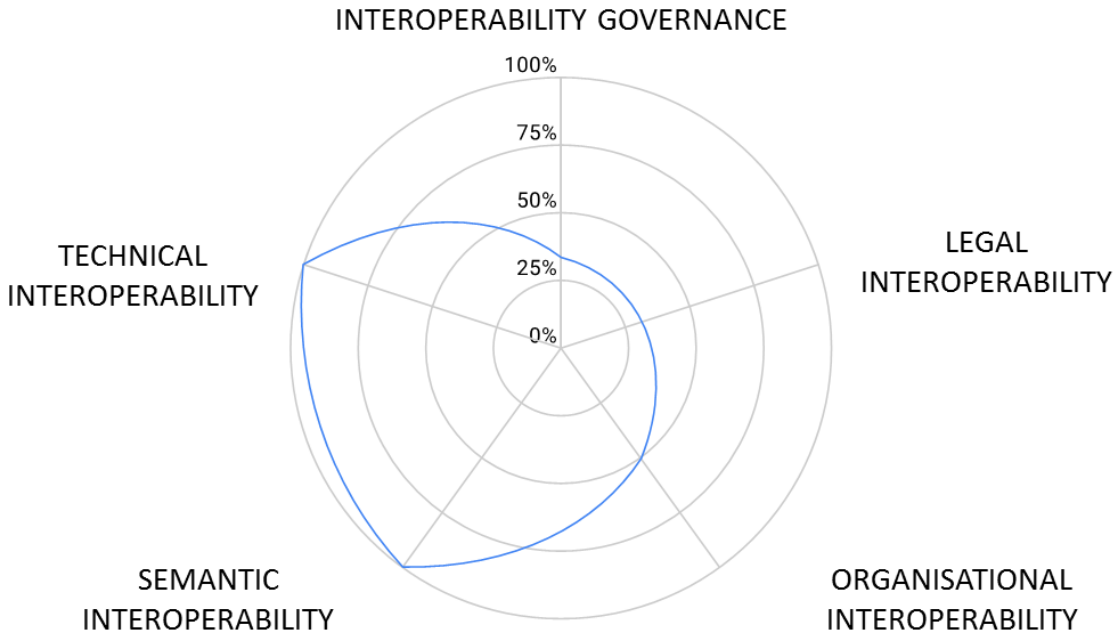
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.1.2 Interpretation according to compliance with EIF

The overall compliance of Beacon with EIF is intermediate.

Out of the four core interoperability principles only transparency achieves complete compliance, whereas openness and technological neutrality score a little bit lower because of not having sufficient market acceptance yet (A8), not being technology agnostic (A16), and not allowing for a partial implementation (A18). Reusability is the dimension with the lowest score due to the fact that the specification is judged not to be available for implementation across business domains (A15) although there are recognizable efforts to develop the specifications to be usable for other domains within the biomedical sciences.

Within the principles related to generic user needs and expectations, the criteria concerning user-centricity, as well as inclusion and accessibility, are entirely met. However, security and privacy only partly complies with EIF since the specification is assessed not to foster a secure and trustworthy data exchange (A25) relying on the systems and context of the implementation for that purpose. The criterion associated with multilingualism is not met (A27) as all the specifications of the standard are in English and, although the standard allows the use of diverse classifications it is not extensible to enable translations of language attributes.

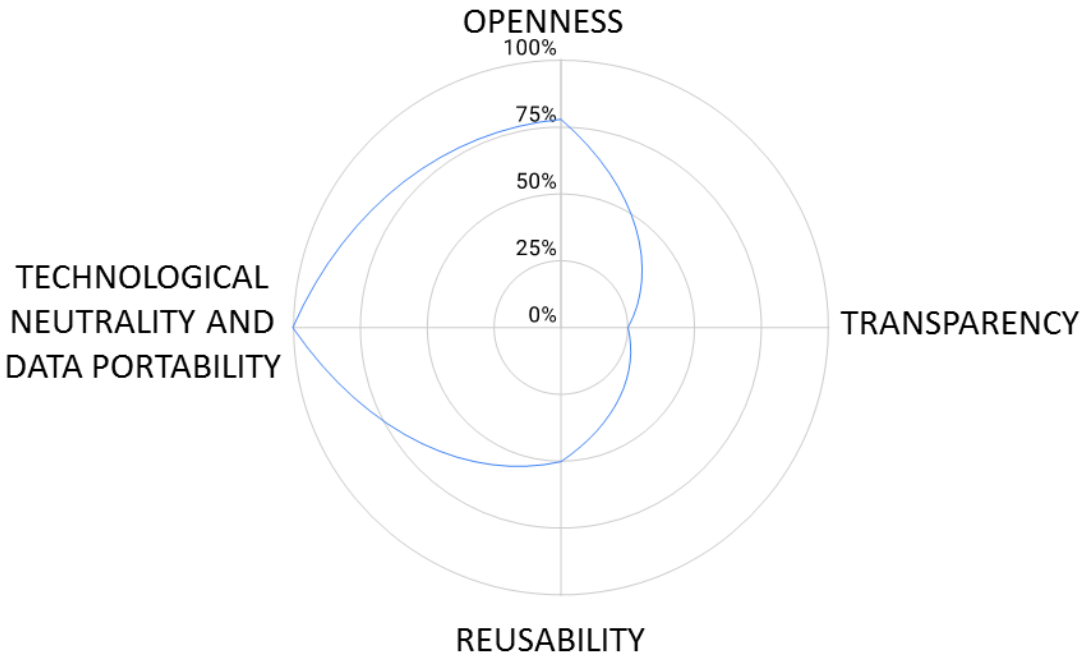
Regarding the foundation principles for cooperation among public administrations, a low score was observed for most criteria. Only administrative simplification was assessed to comply with EIF (A28 and A29). In relation to the assessment of effectiveness and efficiency, it should be remarked that the working group in WP6 was not able to find any documentation supporting the effectiveness and efficiency of the standard as specified by CAMSS. However, this doesn't necessarily mean that it doesn't exist.

The interoperability layers received a high score in the evaluation process for technical interoperability and semantic interoperability, but only an intermediate score for organisational interoperability since it was judged that the standard does not facilitate the modelling of business processes (A40). Interoperability governance was seen only to a small extent to comply with EIF since the standard is not recommended by an EU member state (A35) and not either included in a repository/catalogue of standards at the national level nor at EU level (A37 and A38).

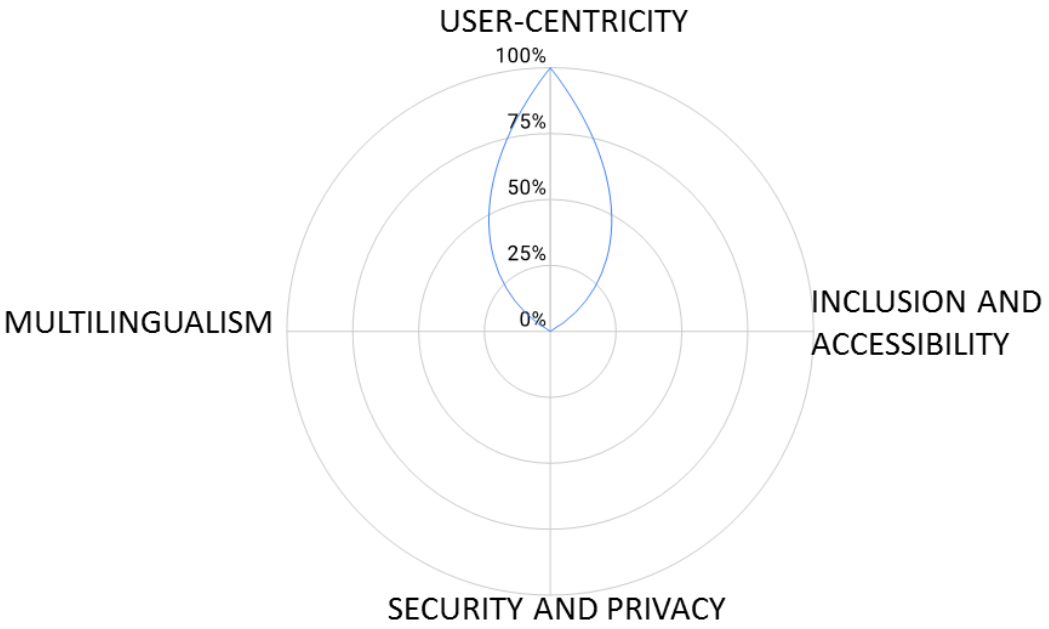
1.1.2 PHIRI-HIP

1.1.2.1 Results in graphs (in percentages)

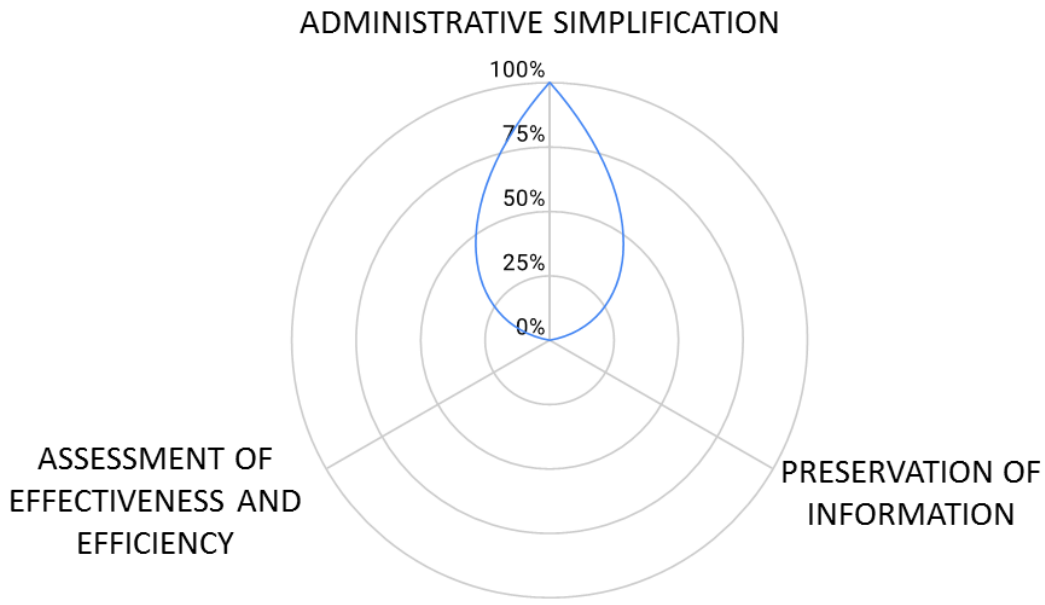
Core interoperability principles



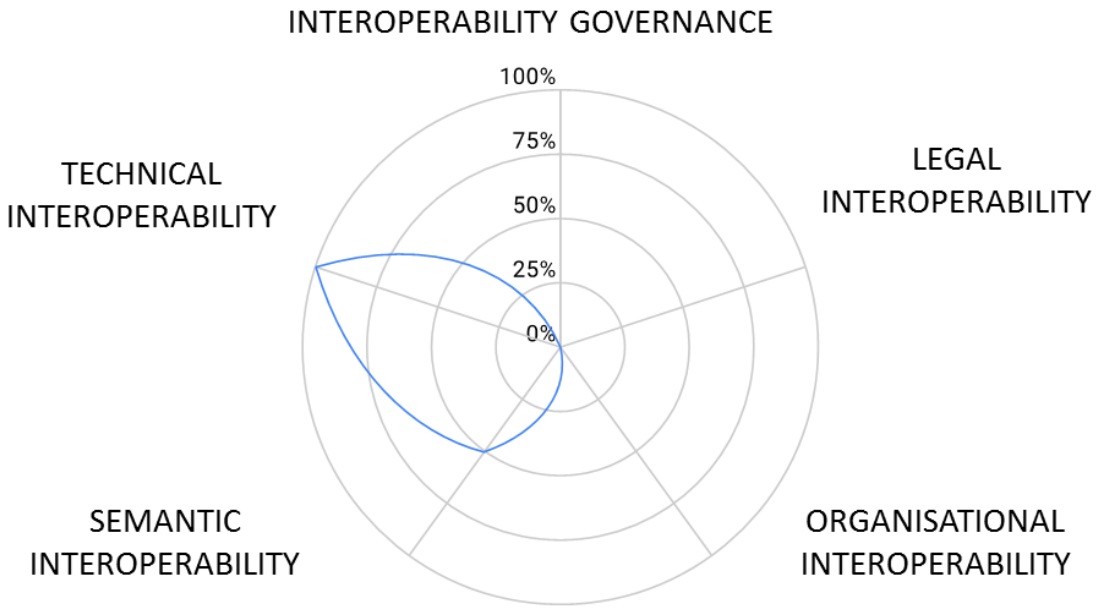
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.2.2 Interpretation according to compliance with EIF

The overall compliance of PHIRI-HIP with EIF was quite low. However, it is to be mentioned that PHIRI-HIP is a fairly new specification with pragmatic application objectives and the PHIRI project is still continuing. The criteria that were compliant with EIF were technological neutrality and data portability, user-centricity, administrative simplification and technical interoperability.

The interoperability area that scored highest was core interoperability principles. Within that area technological neutrality and data portability were completely compliant with EIF. Also, openness scored quite highly. However, the lack of a public review as part of the release lifecycle (A4), as well as our assessment that there yet is not sufficient market acceptance (A8) because the specification is quite new, lowered the score of openness a bit. The transparency criterion was affected because the specification doesn't foster the visibility or scope the comprehensibility of administrative rules, processes, data, services and decision-making of a public administration (A10, A11). Also, the sub-criterion of ensuring the protection of personal data (A13) was not relevant in the case of this specification. Reusability was assessed a bit lower because the specification has not been made available for its implementation across business domains (A15).

Within the other important criteria, the lack of facilitating the modelling of business processes (A40) and organisational interoperability aspects (A41) caused the low score in organisational

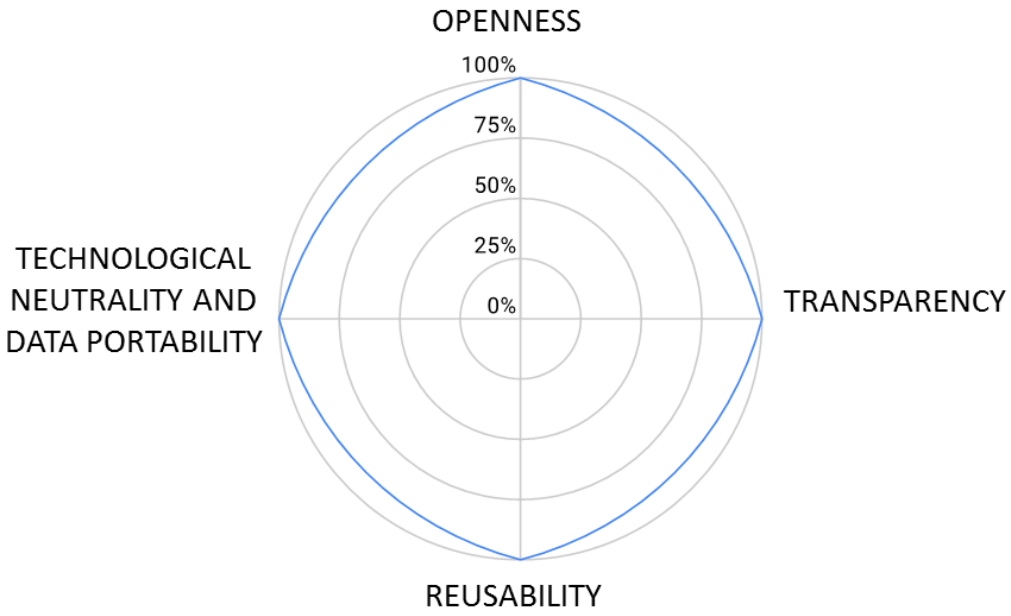
interoperability. The semantic interoperability score was affected because the specification doesn't foster the publication of data as linked open data (A43).

The security and privacy criterion scored 0 because those questions were not relevant in the case of this standard. Also, it is fair to mention that taking into account that PHIRI-HIP is a new specification, it is understandable that assessments of effectiveness and efficiency have not yet been made (A31, A32).

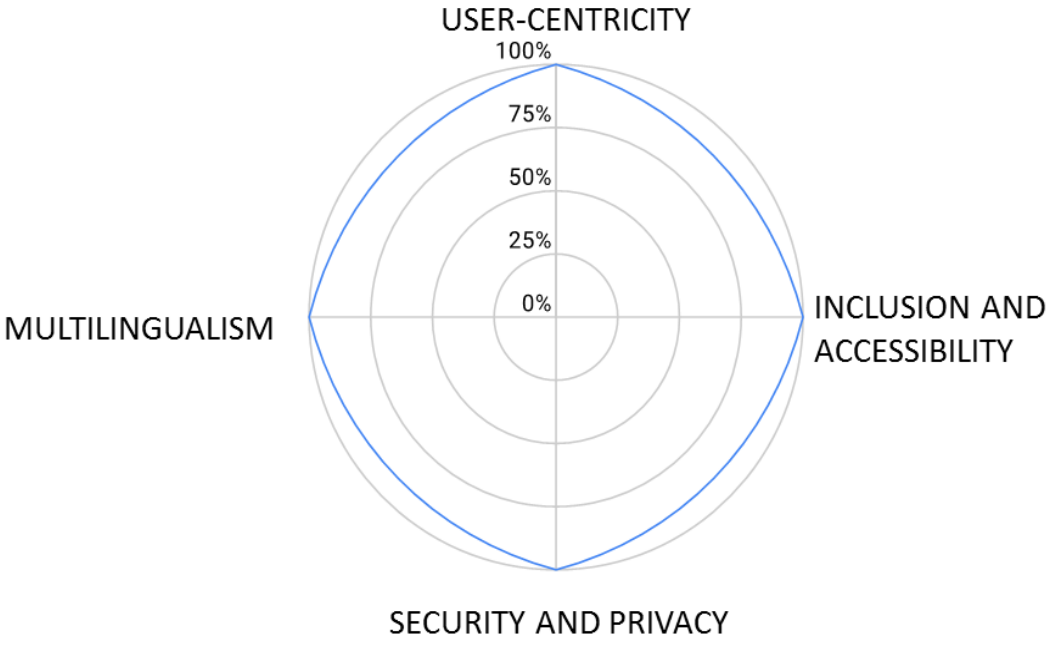
1.1.3 INSPIRE

1.1.3.1 Results in graphs (in percentages)

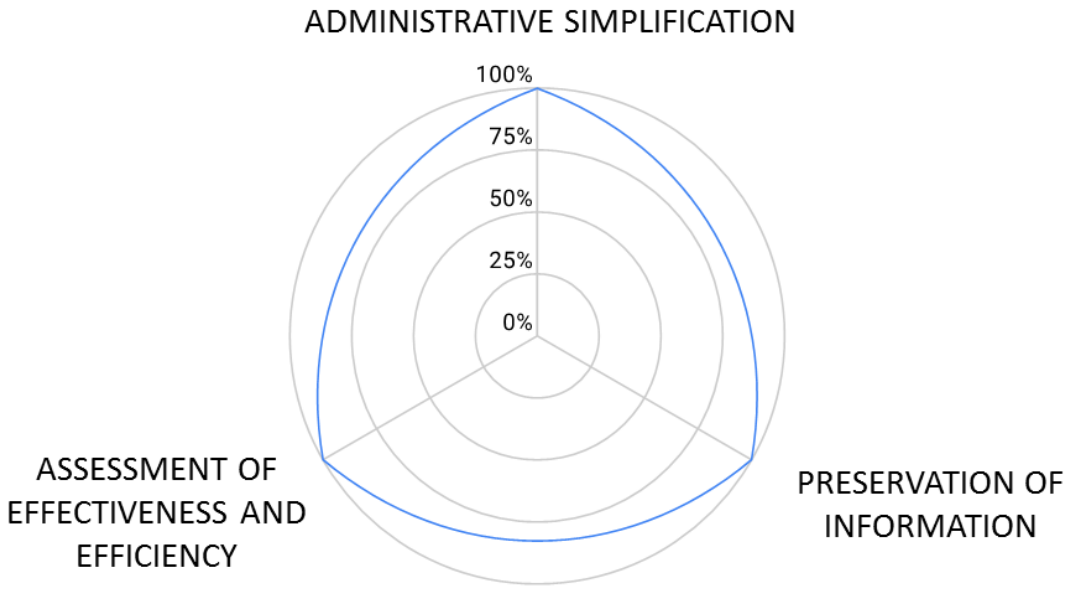
Core interoperability principles



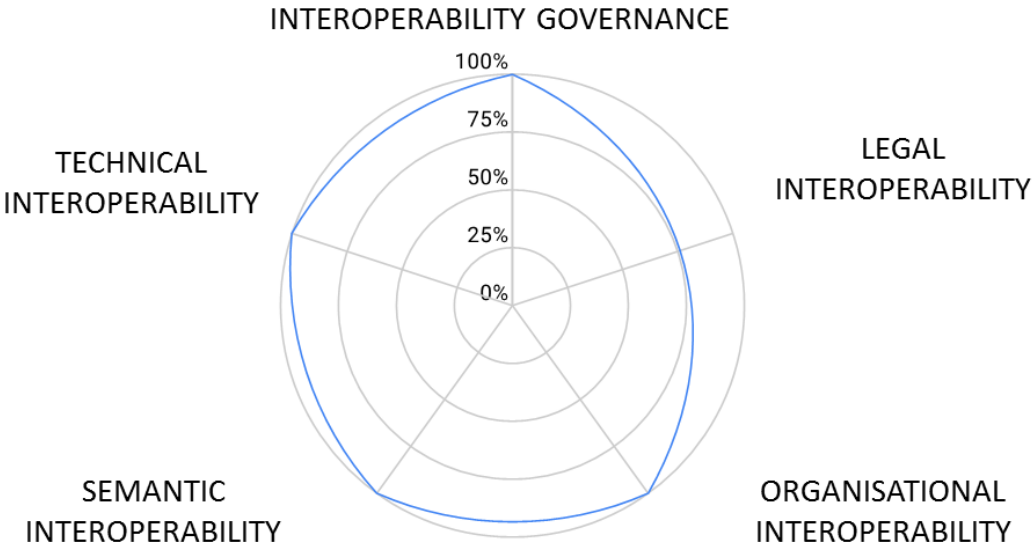
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



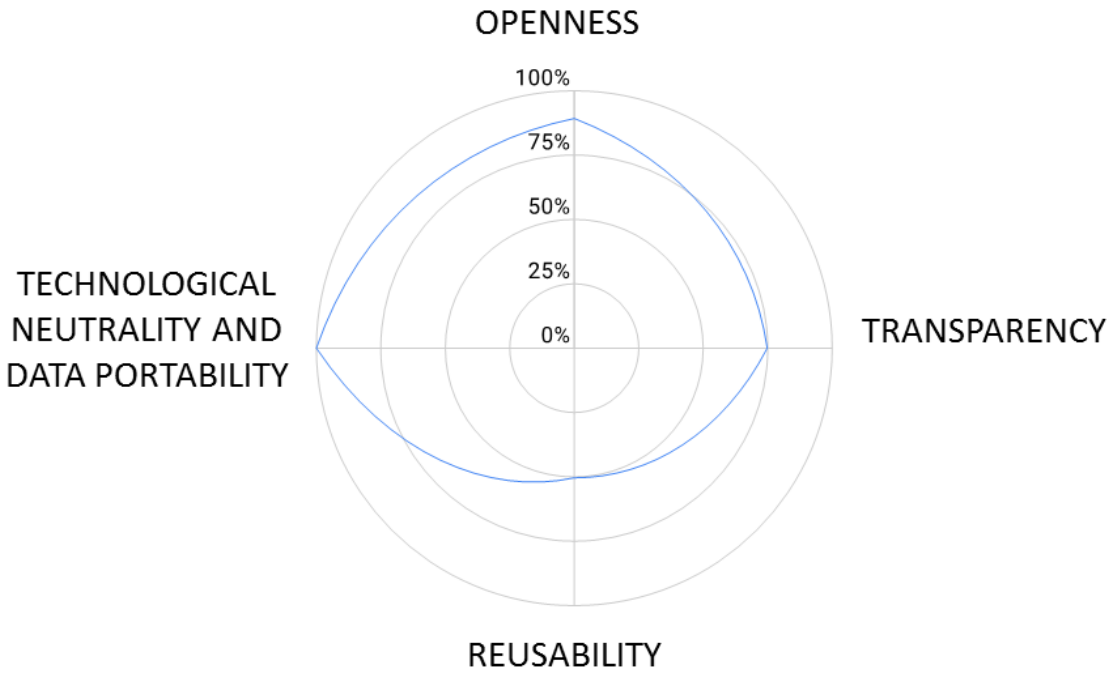
1.1.3.2 Interpretation according to compliance with EIF

INSPIRE was found to be 100 % compliant with EIF for the criteria considered. It should be noted that INSPIRE is an EU standard developed and set for implementation as an EU Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007, establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), published in the official Journal on the 25th April 2007. The INSPIRE Directive entered into force on the 15th May 2007 to ensure that the spatial data infrastructures of the Member States are compatible and usable in a Community and transboundary context, the Directive requires that common Implementing Rules (IR) are adopted in a number of specific areas (Metadata, Data Specifications, Network Services, Data and Service Sharing and Monitoring and Reporting).

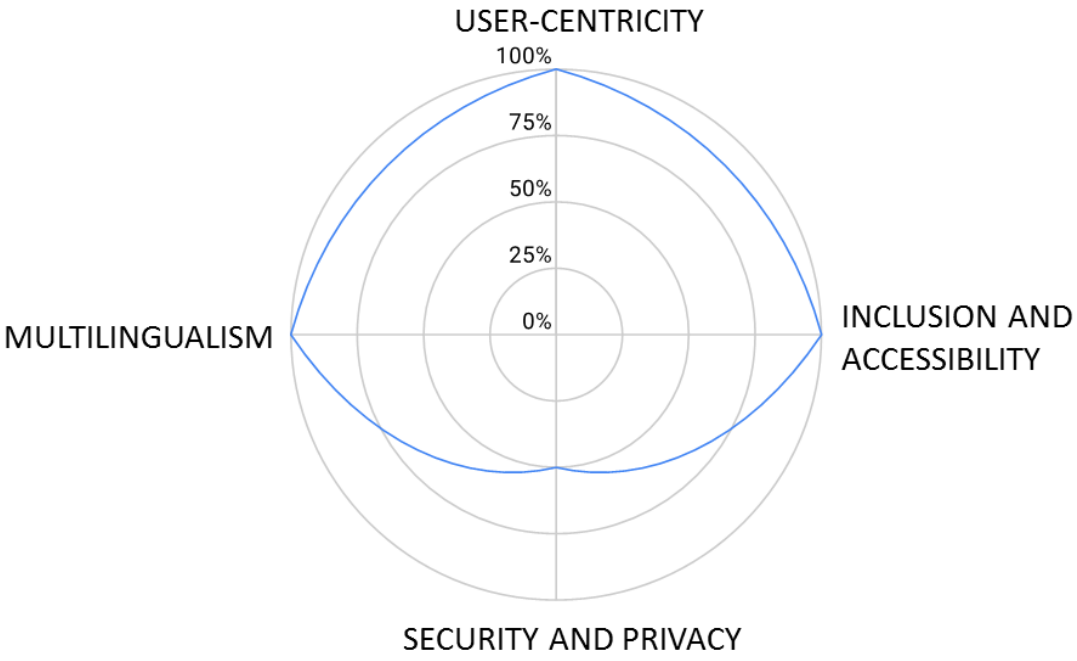
1.1.4 ECRIN CRMDR

1.1.4.1 Results in graphs (in percentages)

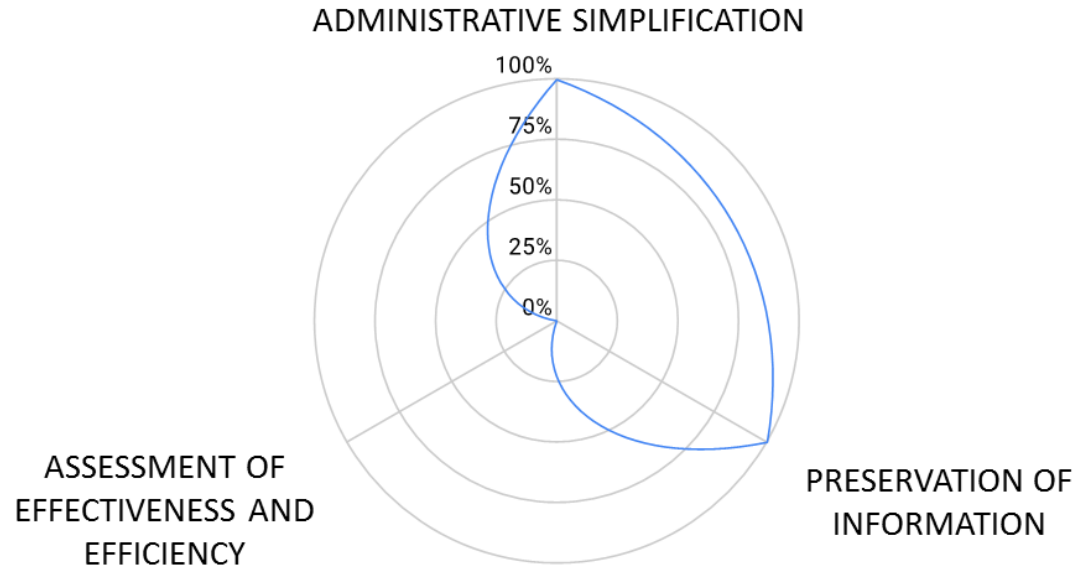
Core interoperability principles



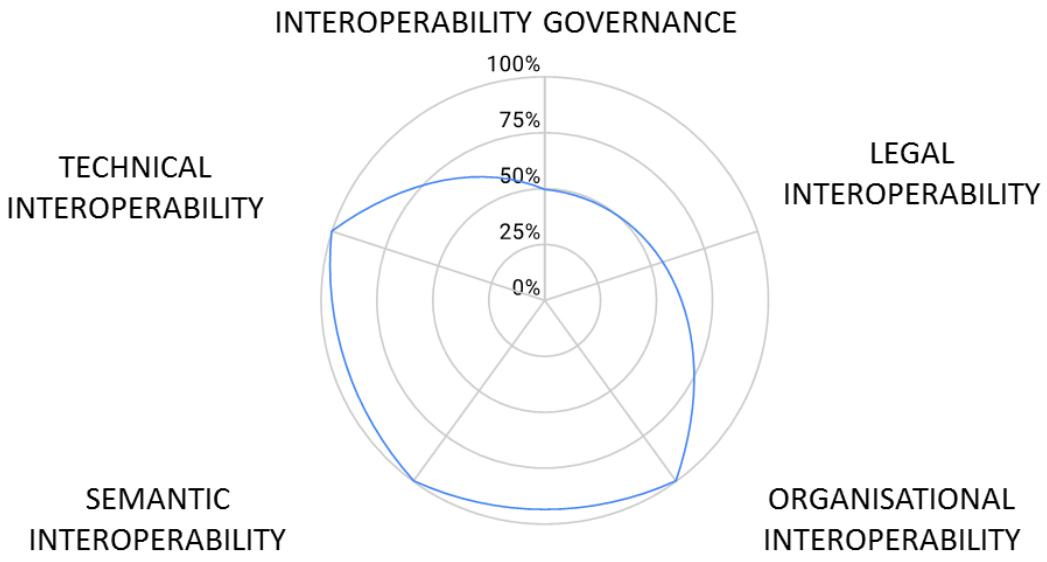
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.4.2 Interpretation according to compliance with EIF

The overall compliance of ECRIN CRMDR with EIF is quite high. The strengths are in the principles related to generic user needs and expectations, as well as in the interoperability layers.

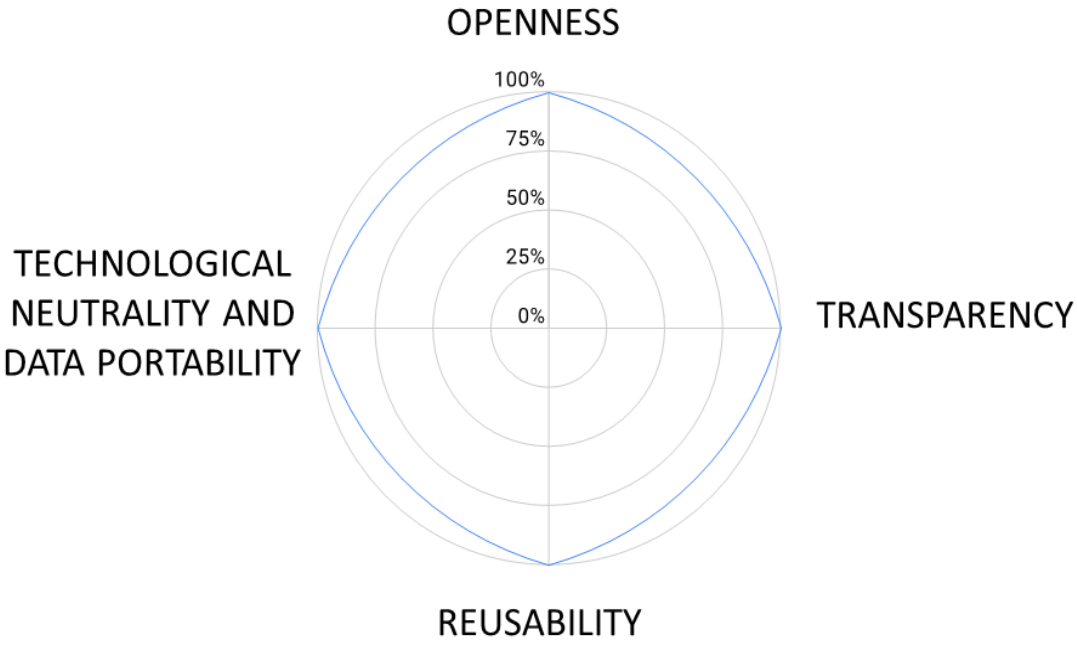
The area that scored lowest was foundation principles for cooperation among public administrations because assessments of effectiveness and efficiency could not be found (A31, A32).

In the other areas, the openness score was lowered a bit because a public review is not part of the current release lifecycle (A4) even though a public review would be possible. The transparency score was affected because the standard does not ensure the protection of personal data (A13) since personal data does not fall within the schema. Reusability was lowered because it was assessed that the standard has not been made available for implementation across business domains (A15).

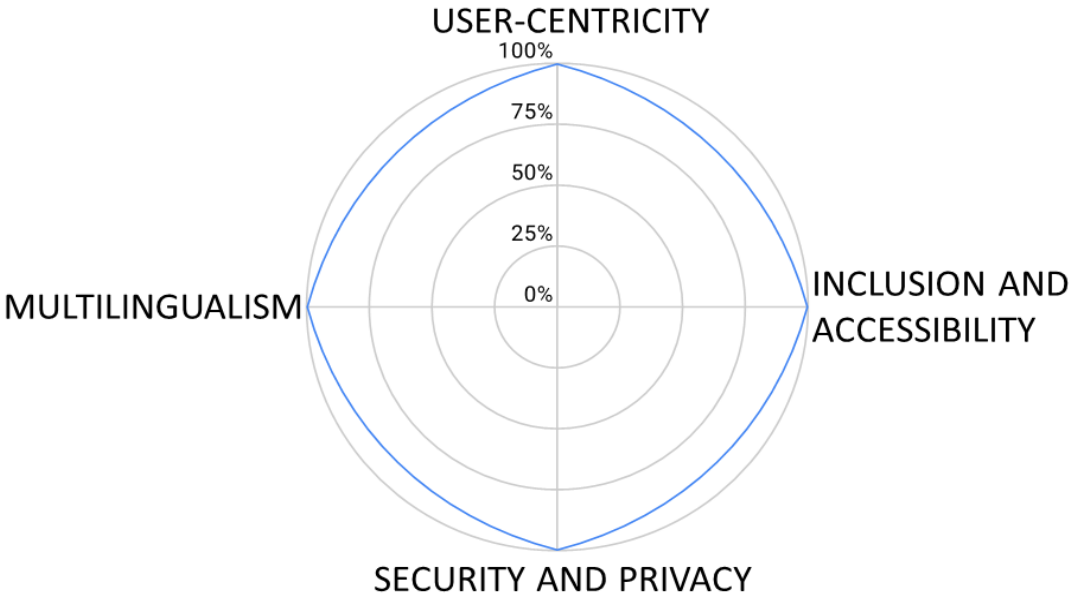
1.1.5 DCAT-AP

1.1.5.1 Results in graphs (in percentages)

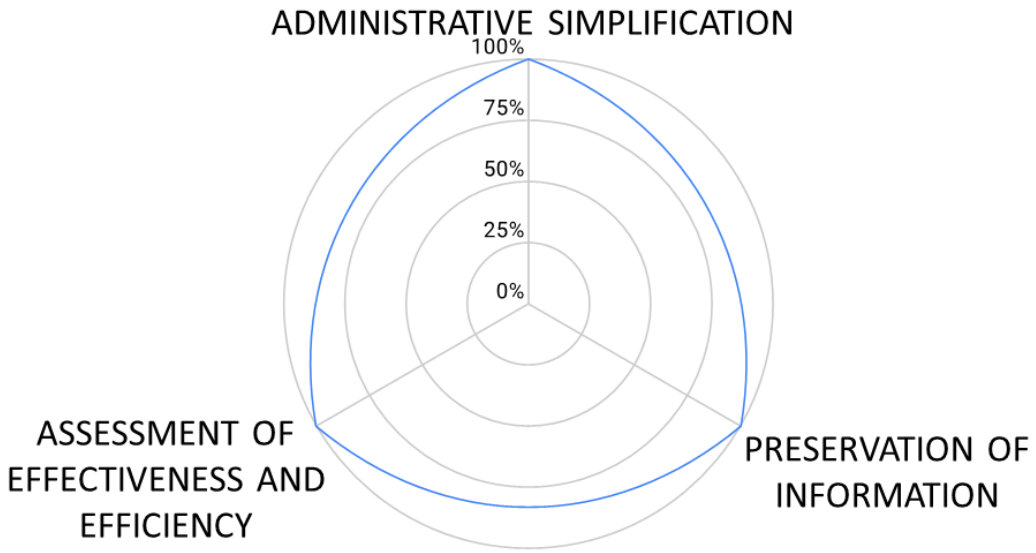
Core interoperability principles



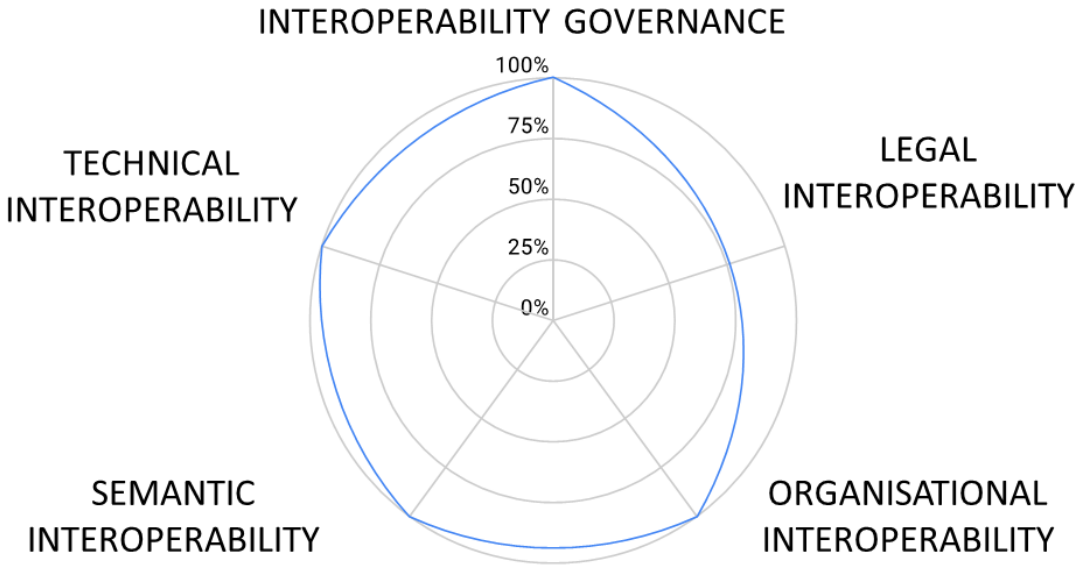
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



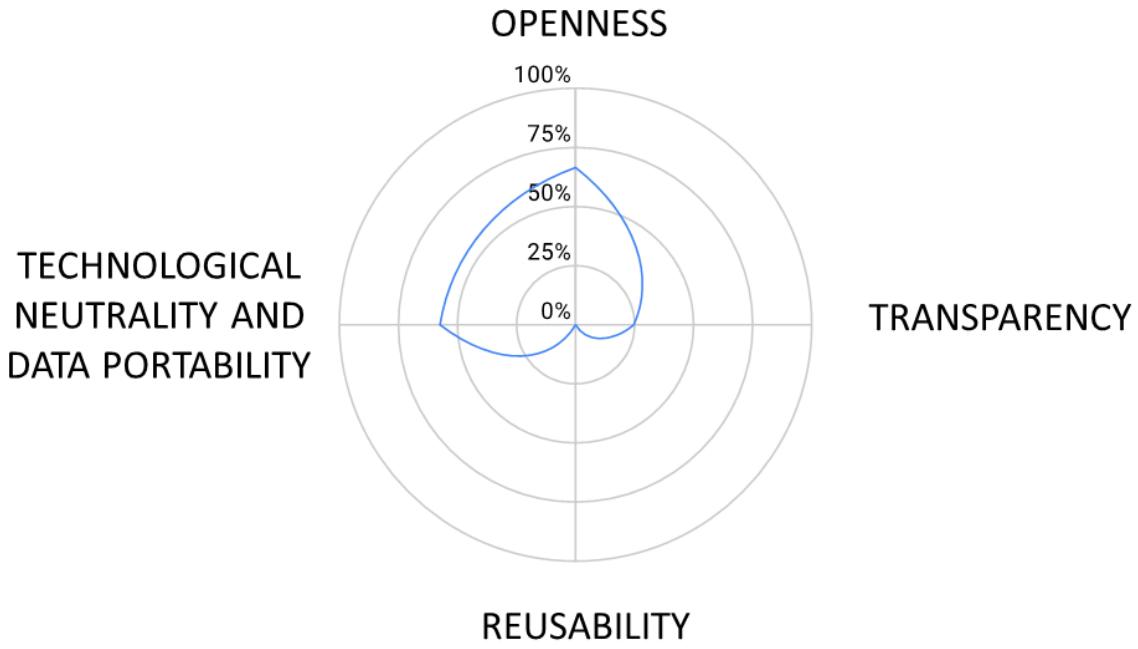
1.1.5.2 Interpretation according to compliance with EIF

DCAT-AP was found to be completely compliant with EIF.

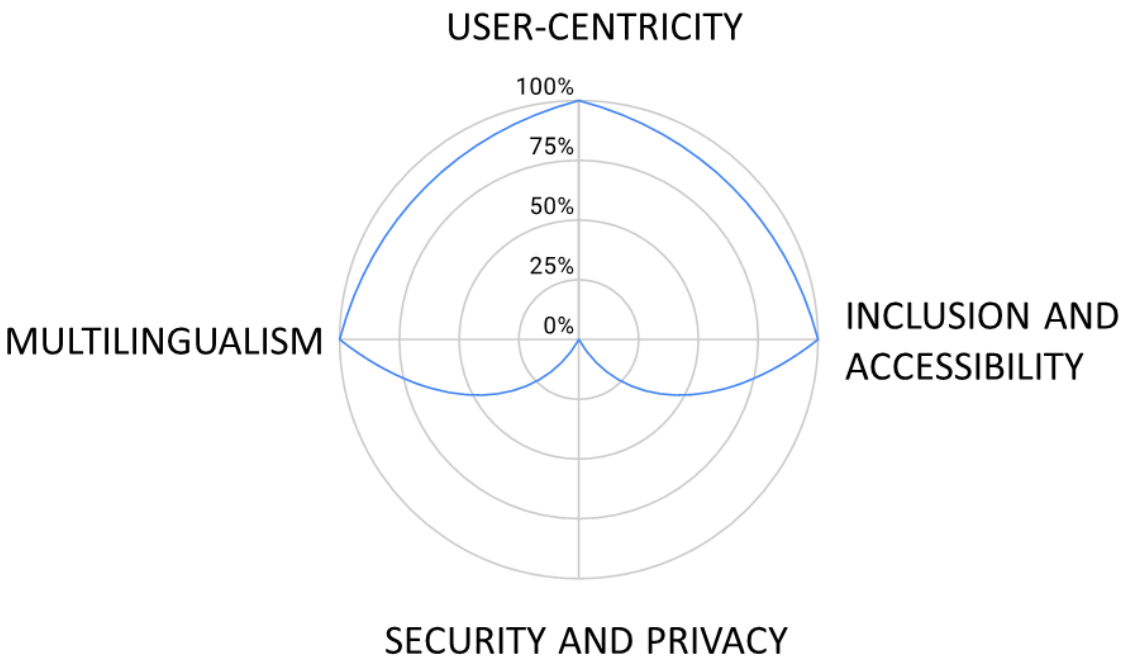
1.1.6 Bioimage archive

1.1.6.1 Results in graphs (in percentages)

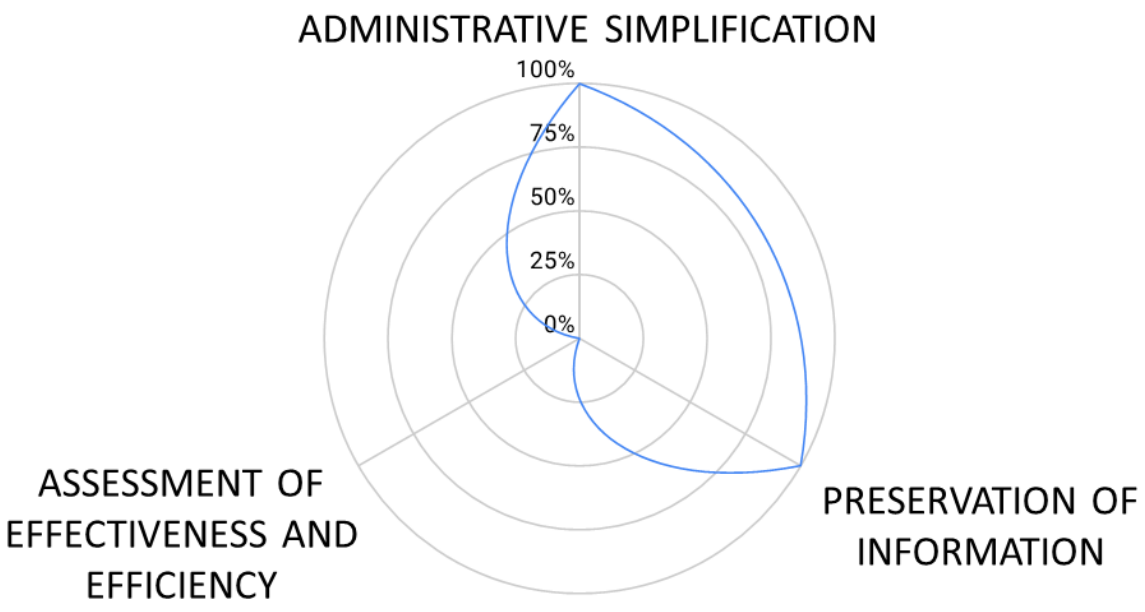
Core interoperability principles



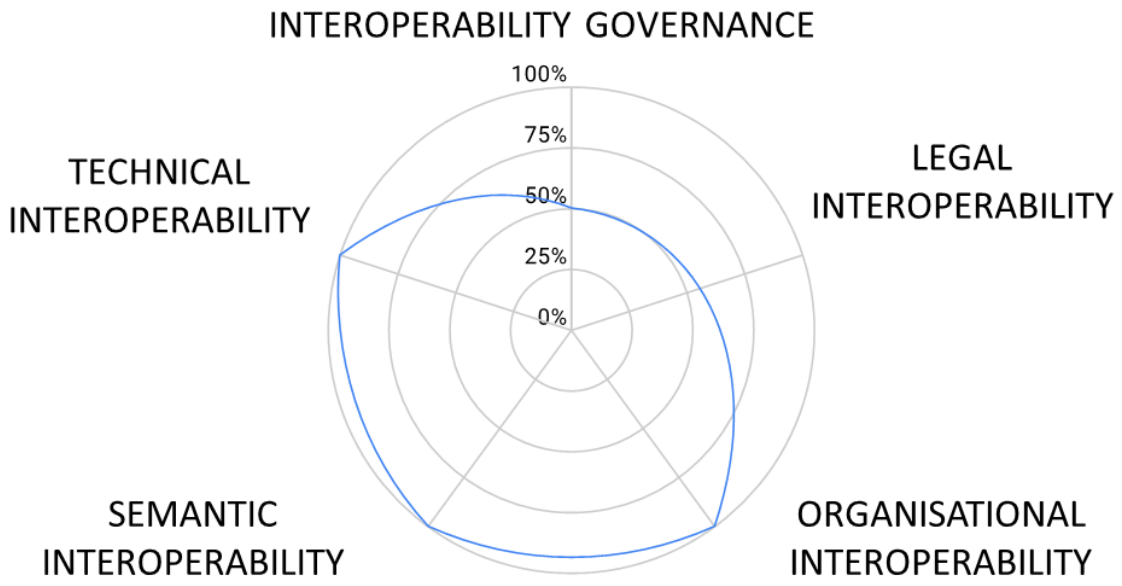
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.6.2 Interpretation according to compliance with EIF

The overall compliance of Bioimage Archive with EIF is not very high.

The core interoperability principles scored fairly low on all parameters - ranging from 0 for reusability to intermediate for openness, as well as technological neutrality and data portability. The evaluation of openness can be attributed to the fact that stakeholders do not have the opportunity to contribute to the development of the specification (A3), a public review is not part of the decision-making process (A4), and the specification does not have sufficient market acceptance for its use in the development of products and services (A8). Transparency is low mainly because it is judged that the specification does not foster the visibility nor scopes the comprehensibility of administrative rules, processes, data, services, and decision-making of a public administration (A10 and A11).

The principles related to generic user needs and expectations are in general met - with exception of the principles of security and privacy (A25 and A26) which are judged not to be relevant for this standard.

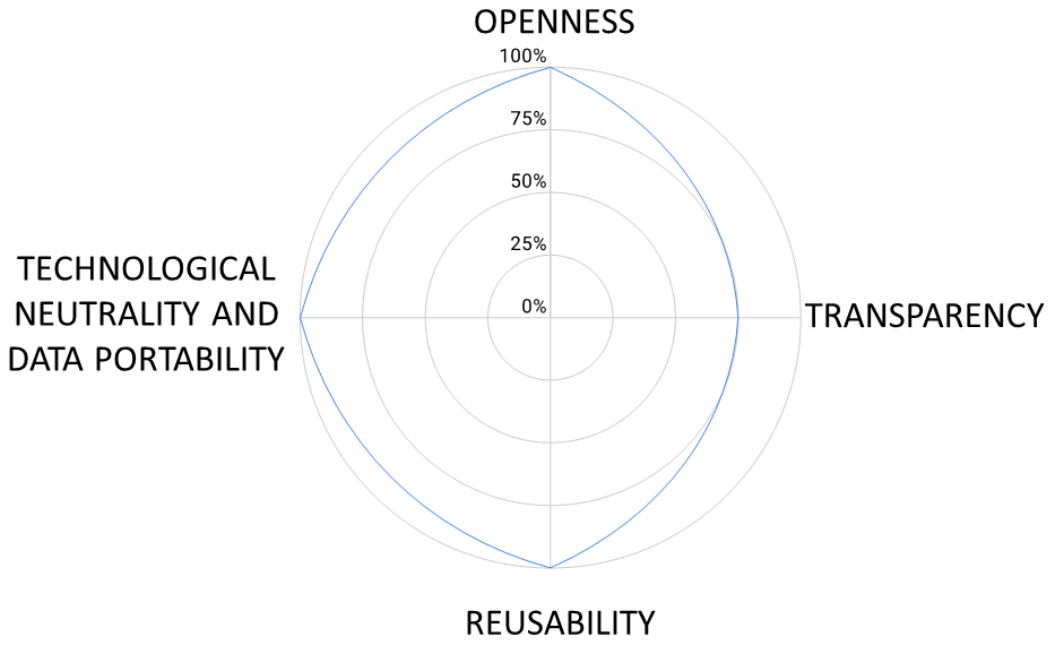
In general, the foundation principles for cooperation among public administrations have a high degree of compliance with EIF. However, it has not been possible to find existing studies nor documentation assessing the standard or specification in terms of effectiveness and efficiency (A31 and A32).

Three of the criteria concerning interoperability layers are fully met, but interoperability governance scores lower since the specification is not recommended by an EU Member State (A35) and not included in a repository/catalogue of standards at the national level (A38), although in a centralised repository. It has not been judged relevant to consider whether the standard can be mapped to the European Interoperability Reference Architecture (A33). The legal interoperability criterion (A39) is not applicable for this standard.

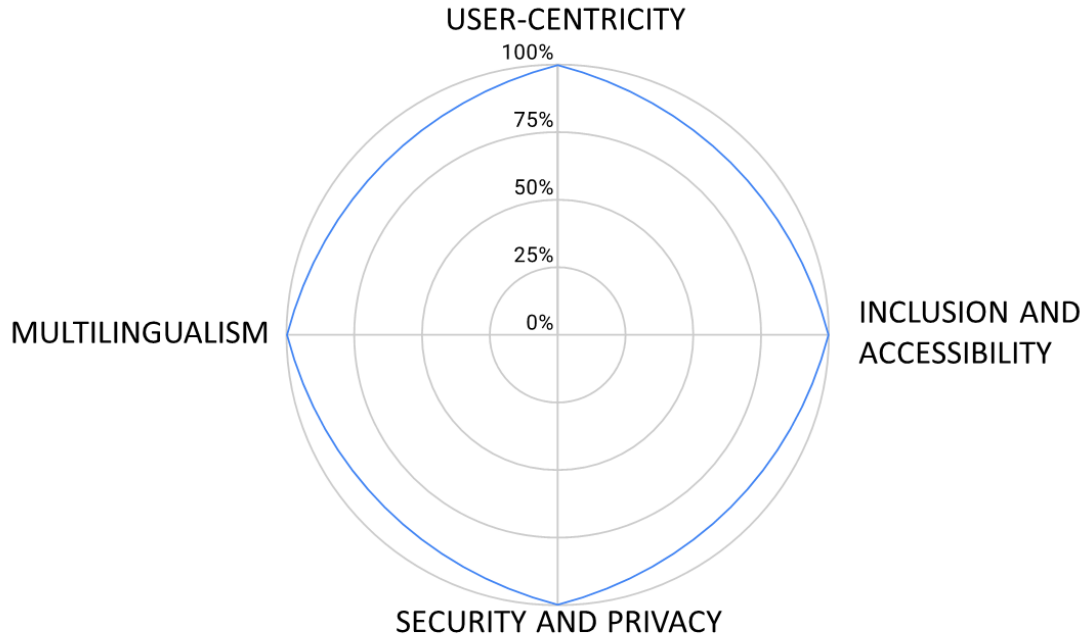
1.1.7 CESSDA CMM

1.1.7.1 Results in graphs (in percentages)

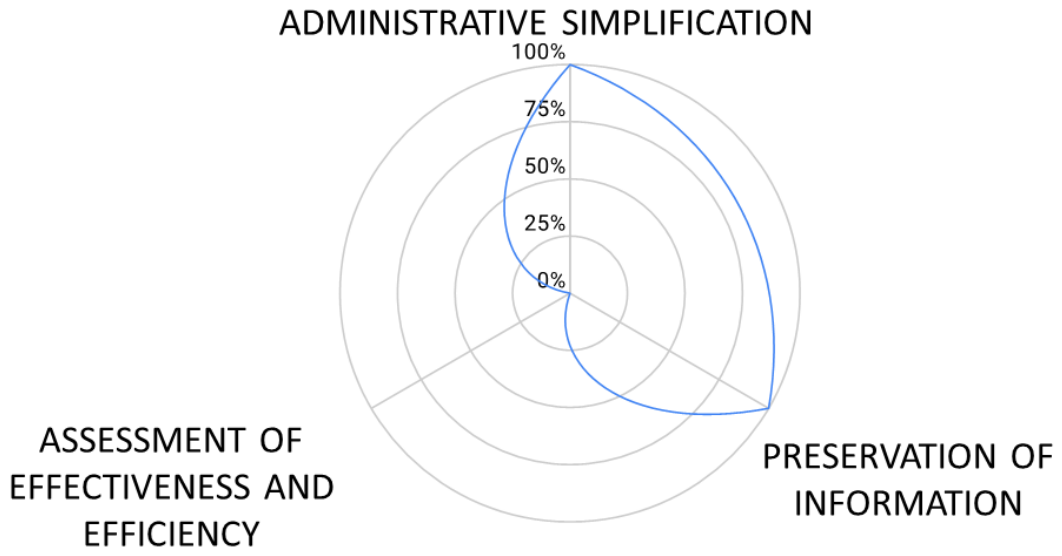
Core interoperability principles



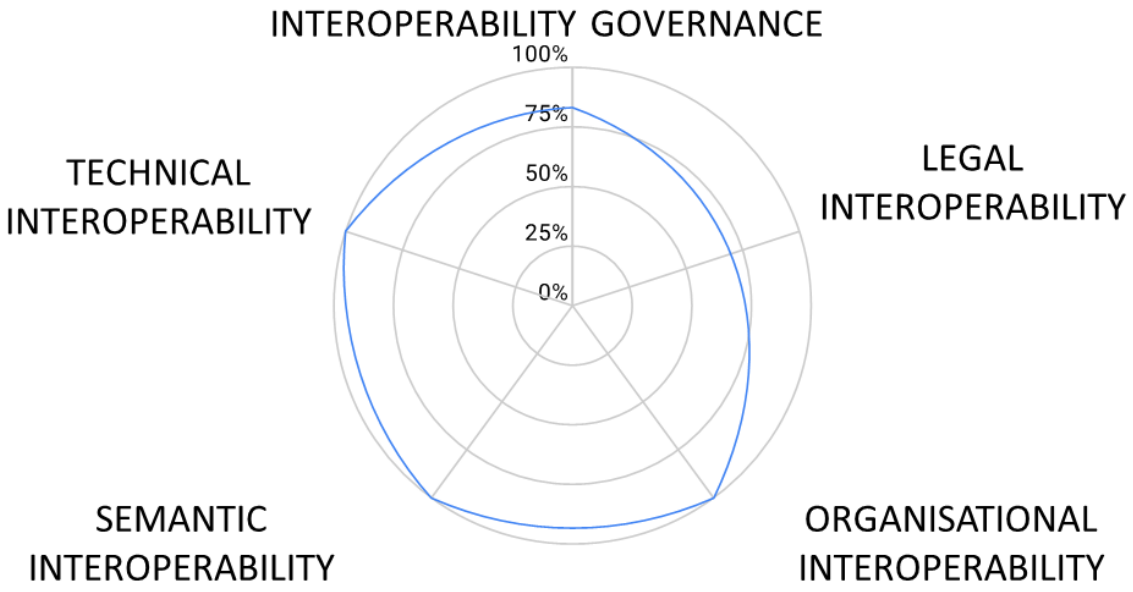
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.7.2 Interpretation according to compliance with EIF

The overall compliance of CESSDA CMM with EIF is high. The strongest area is the principles related to generic user needs and expectations with 100% compliance.

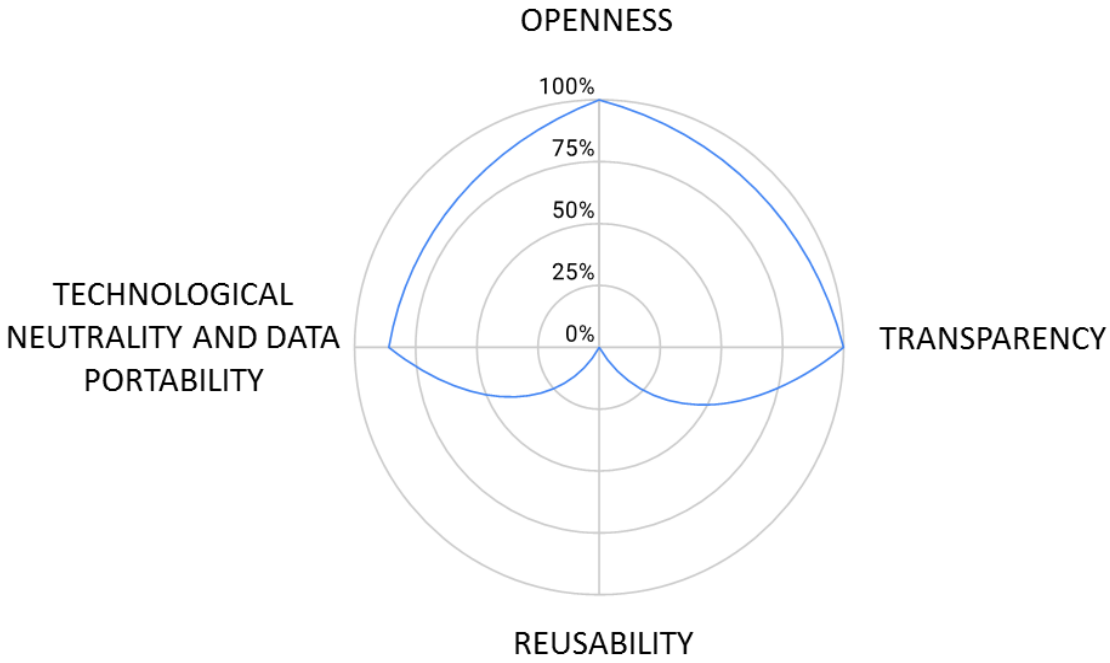
The core interoperability principles score highly, including openness and reusability. Only the transparency criterion has less than 100% compliance within this group. Transparency scored a bit lower because it was assessed that the standard only partly scopes the comprehensibility of administrative rules, data, services, and decision-making of a public administration (A11).

Within the principles related to generic user needs and expectations the criterion assessment of effectiveness and efficiency scored 0% because assessments of effectiveness and efficiency of the standard have not been made (A31, A32). The interoperability layers area is strong with only the interoperability governance criterion scoring lower than 100% because it was assessed that the specification is not recommended by an EU Member State (A35).

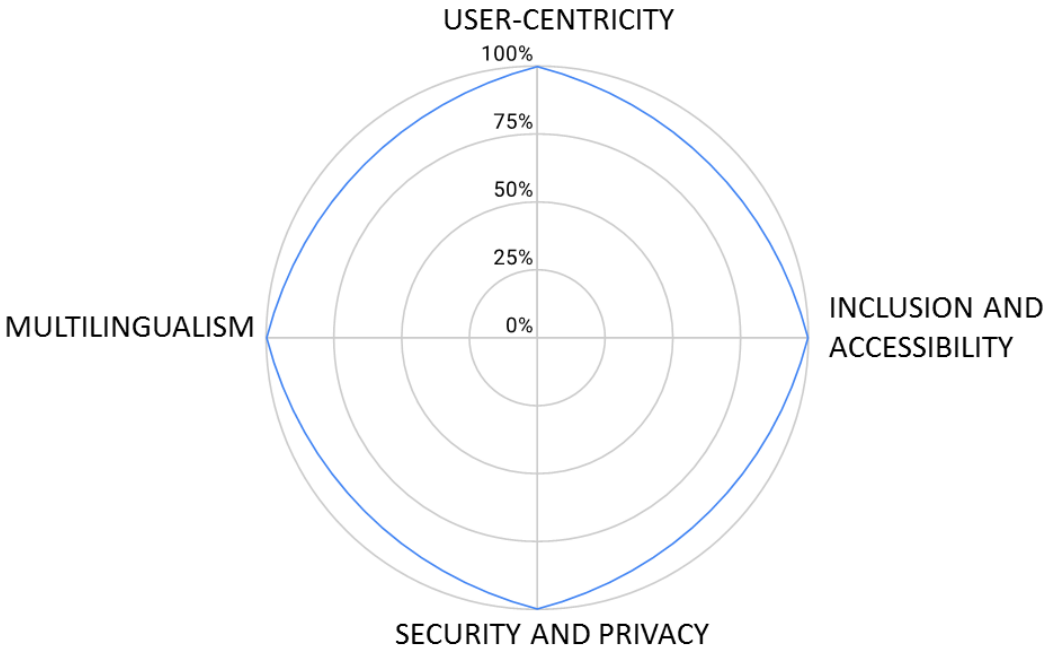
1.1.8 BBMRI MIABIS

1.1.8.1 Results in graphs (in percentages)

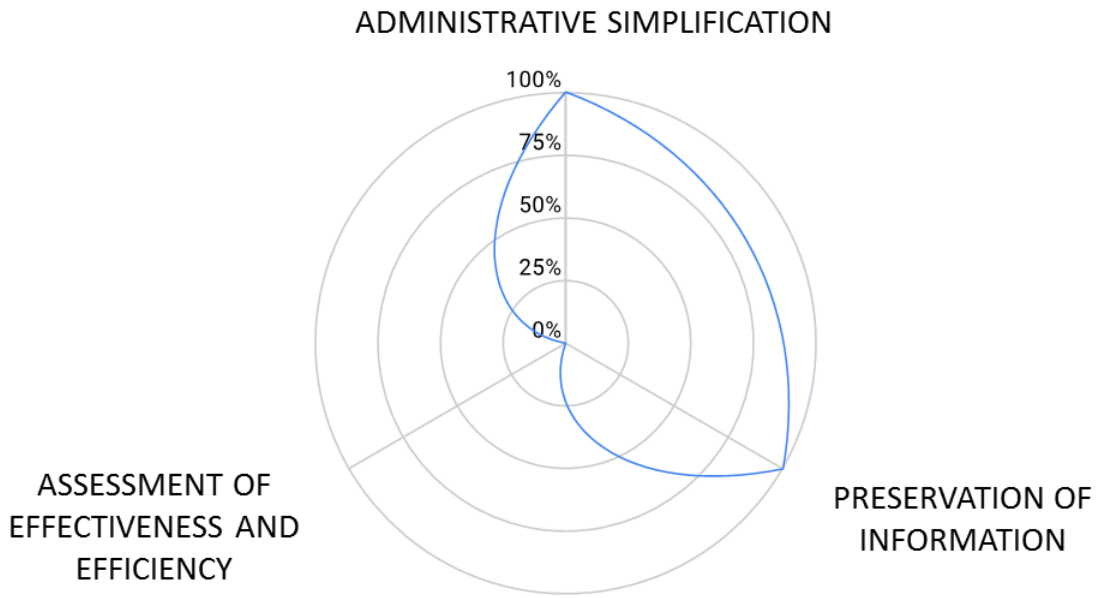
Core interoperability principles



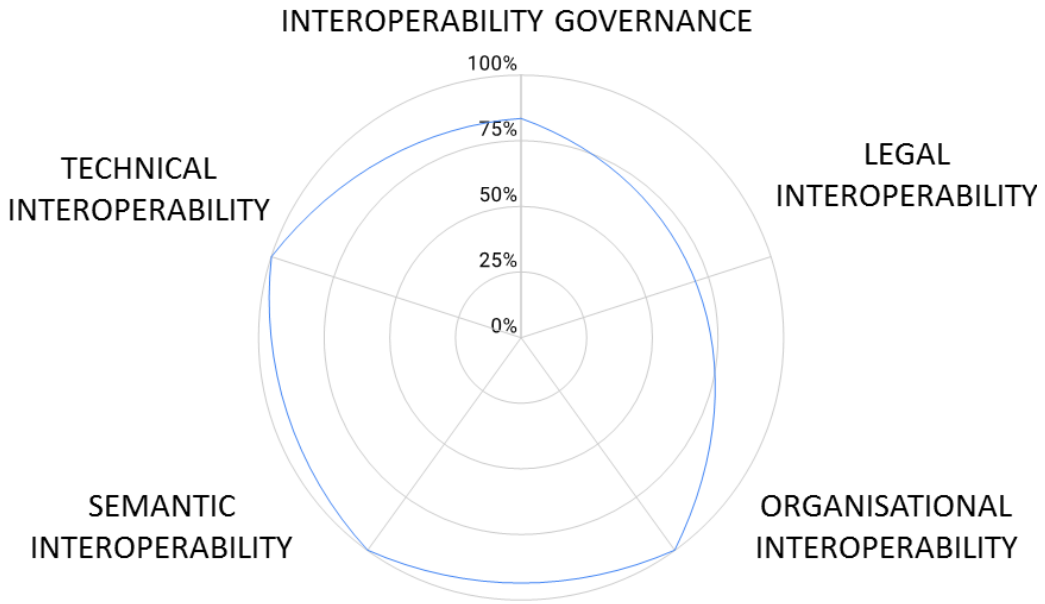
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.8.2 Interpretation according to compliance with EIF

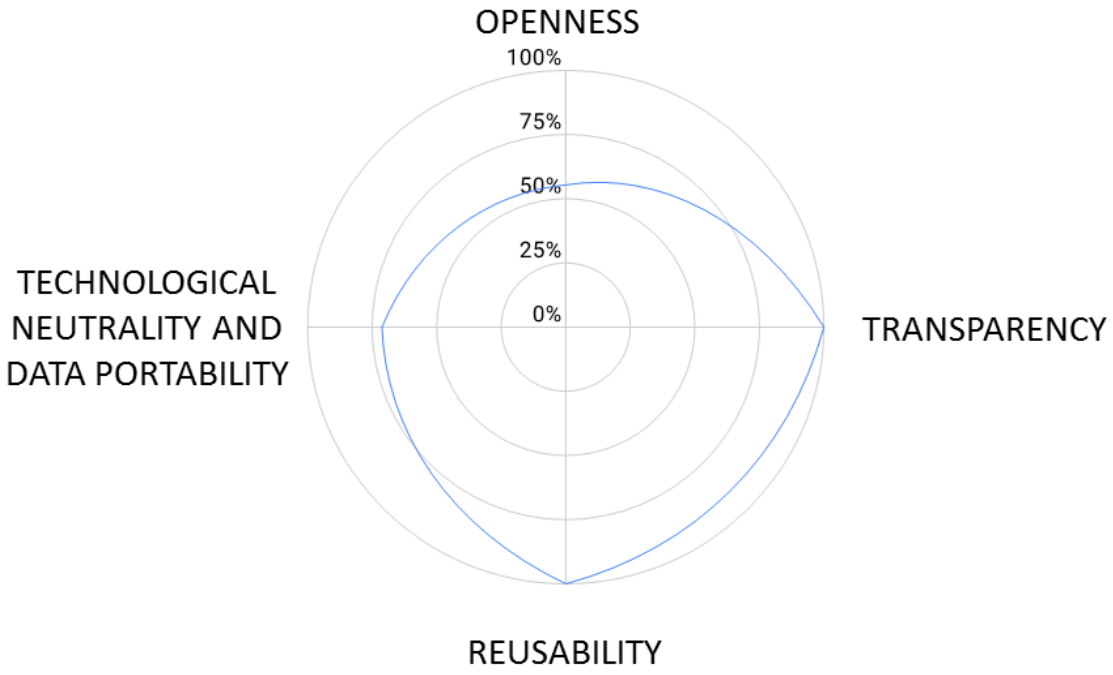
The overall compliance of BBMRI MIABIS with EIF is high.

The few weaknesses were found in reusability, assessment of effectiveness and efficiency and to some extent in technological neutrality and data portability.
Reusability scored low because it was assessed that the standard is not reusable across business domains (A14, A15). Assessment of effectiveness and efficiency scored 0 because studies assessing effectiveness or efficiency could not be found (A31, A32). Also the technological neutrality and data portability criterion was lowered a bit because it was assessed that the specification is not platform agnostic (A17). The interoperability governance criterion was lowered because the standard is not mapped to the European Interoperability Reference Architecture EIRA (A33).

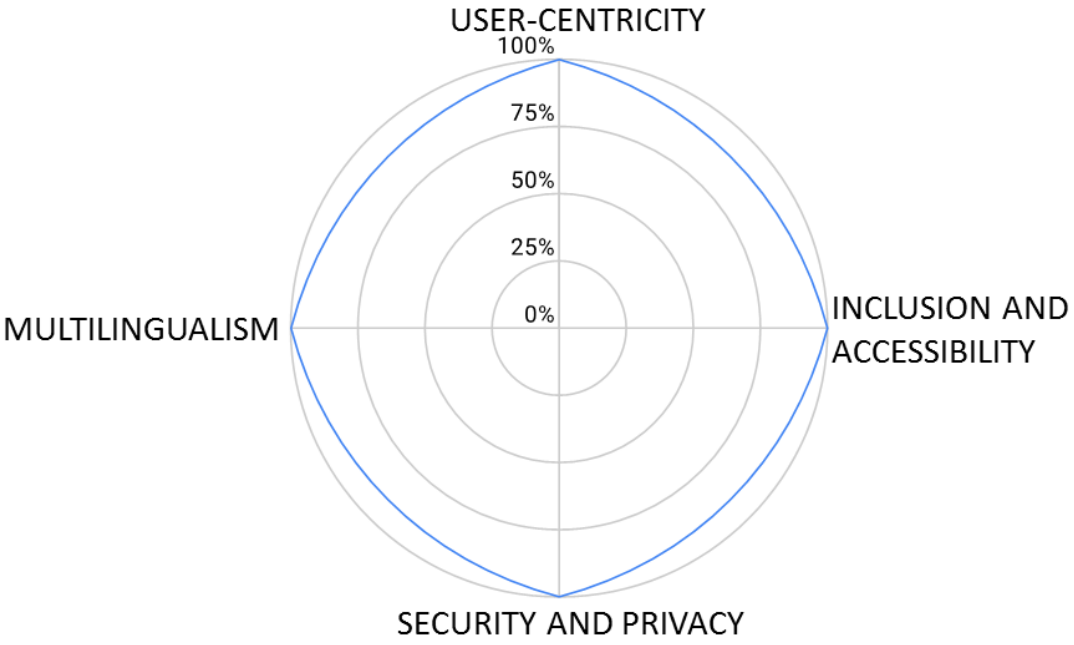
1.1.9 FairSharing

1.1.9.1 Results in graphs (in percentages)

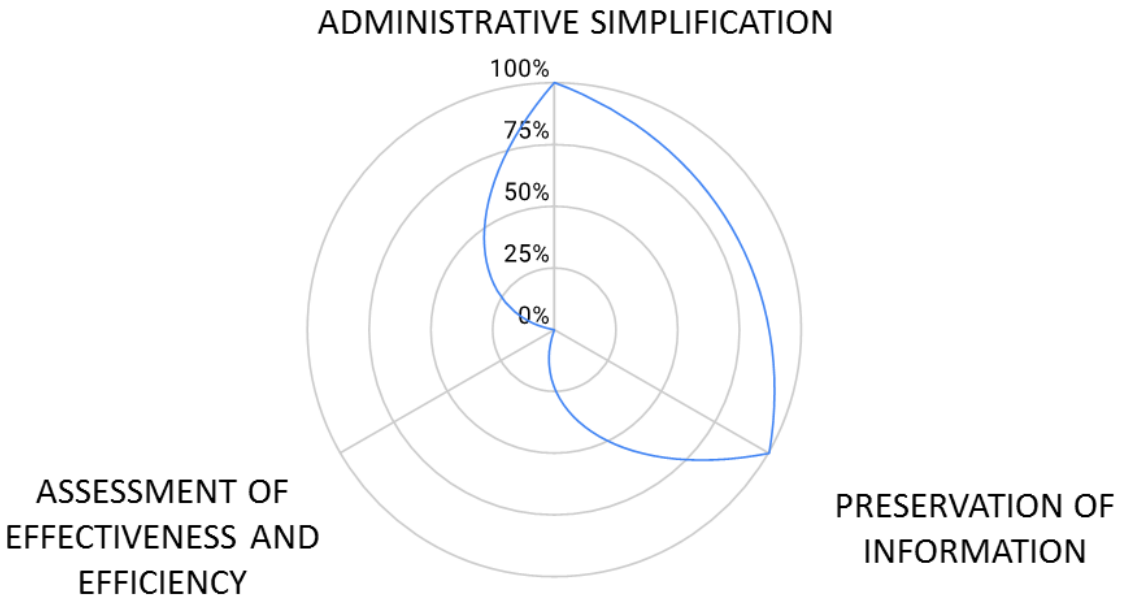
Core interoperability principles



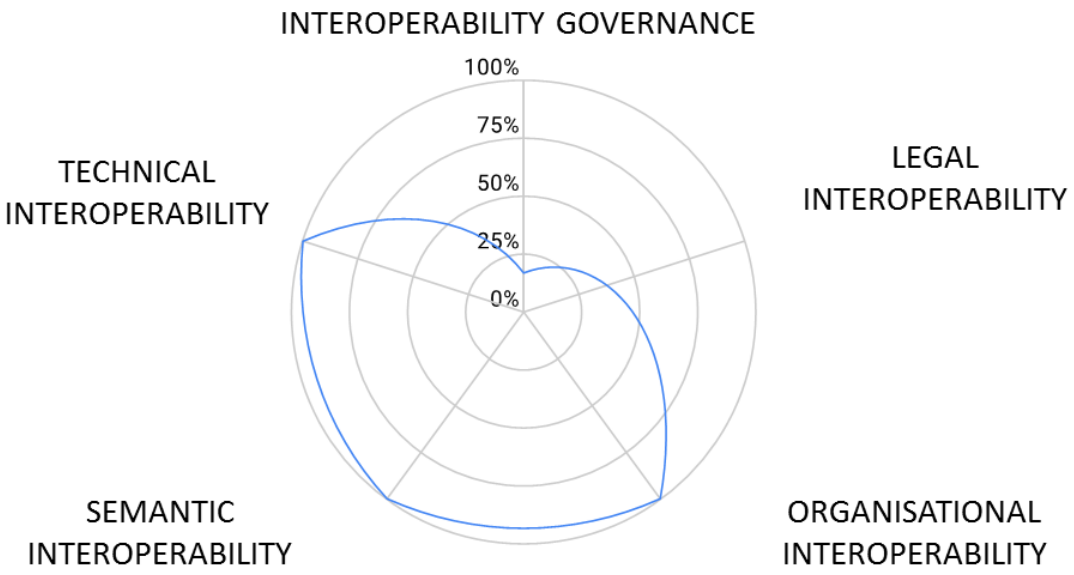
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.1.9.2 Interpretation according to compliance with EIF

Fairsharing was found to have a relatively high degree of compliance with EIF.

Within the core interoperability principles, reusability and transparency fully comply. Openness and technological neutrality and data portability both score at an intermediate level since the standard is not judged to be sufficiently mature (A7) and have a sufficient market acceptance (A8) with respect to openness, and since it does not allow customisation and extension (A19 and A20) with respect to technological neutrality and data portability.

The criteria within the principles related to generic user needs and expectations are fully met.

Regarding the foundation principles for cooperation among public administrations, the criteria associated with administrative simplification and preservation of information are completely fulfilled. However, no documentation of assessment of effectiveness and efficiency has been found.

In the interoperability layers, technical interoperability, semantic interoperability, and organisational interoperability fully meet the criteria. For the interoperability governance dimension, the evaluation score is low since the standard is not recommended by an EU member state, has not been selected for use in an EU border cross project, and is not included in a repository/catalogue of standards at the national level nor at EU level (A35 - A38). Finally, the criterion concerning legal interoperability is not applicable (A39).

1.2 Standards enabling semantic interoperability

In the semantic standard group, there are standards or specifications which aim to provide a semantic layer for interoperability. Under this role there are taxonomies, ontologies and common data models.

The standards in this group are Orphanet standards, OMOP CDM, CDISC SDTM, SNOMED CT and LOINC. The average scores for each are shown in table 2.

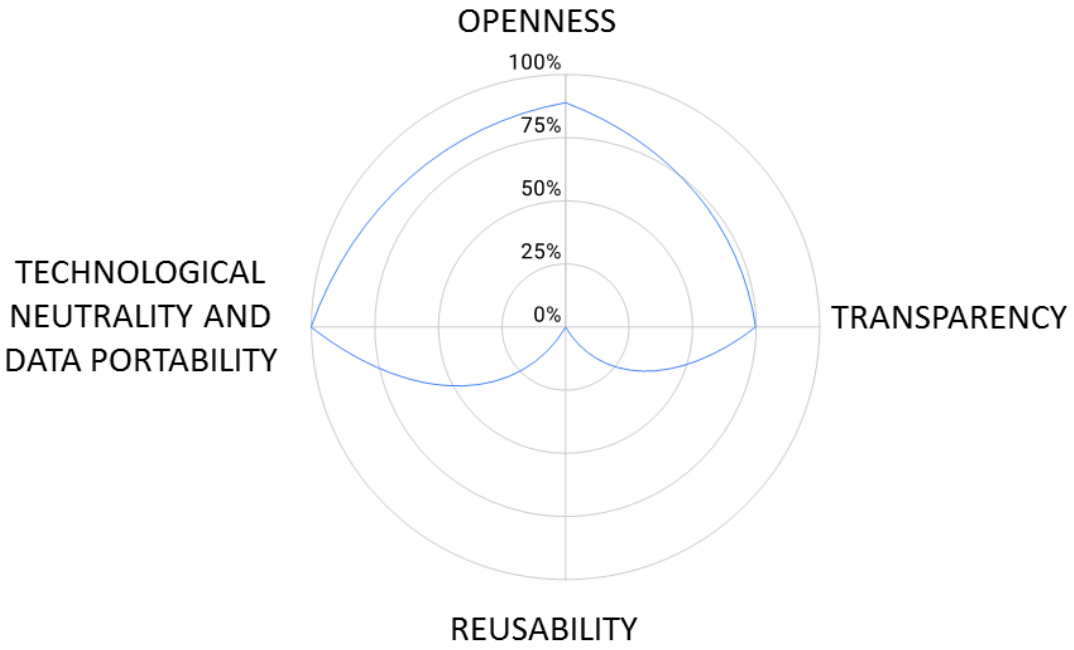
Table 2. Average scores for semantic standards

CRITERION	CDISC/SDTM	LOINC	OMOP-CDM	Orphanet standards	SNOMED-CT
CORE INTEROPERABILITY PRINCIPLES	69%	75%	71%	66%	69%
OPENNESS	89%	100%	100%	89%	78%
TRANSPARENCY	100%	100%	100%	75%	100%
REUSABILITY	0%	0%	0%	0%	0%
TECHNOLOGICAL NEUTRALITY AND DATA PORTABILITY	86%	100%	86%	100%	100%
USERS' NEEDS AND EXPECTATIONS	100%	75%	88%	75%	75%
USER-CENTRICITY	100%	100%	100%	100%	100%
INCLUSION AND ACCESSIBILITY	100%	100%	100%	100%	100%
SECURITY AND PRIVACY	100%	0%	50%	0%	0%
MULTILINGUALISM	100%	100%	100%	100%	100%
PRINCIPLES FOR COOPERATION BETWEEN INSTITUTIONS	67%	83%	100%	67%	83%
ADMINISTRATIVE SIMPLIFICATION	50%	100%	100%	100%	100%
PRESERVATION OF INFORMATION	100%	100%	100%	100%	100%
ASSESSMENT OF EFFECTIVENESS AND EFFICIENCY	50%	50%	100%	0%	50%
INTEROPERABILITY LAYERS	71%	79%	71%	88%	83%
INTEROPERABILITY GOVERNANCE	33%	67%	33%	50%	83%
LEGAL INTEROPERABILITY	N/A	N/A	N/A	N/A	N/A
ORGANISATIONAL INTEROPERABILITY	100%	100%	100%	100%	100%
SEMANTIC INTEROPERABILITY	50%	50%	50%	100%	50%
TECHNICAL INTEROPERABILITY*	100%	100%	100%	100%	100%
OVERALL SCORE	77%	78%	82%	74%	81%

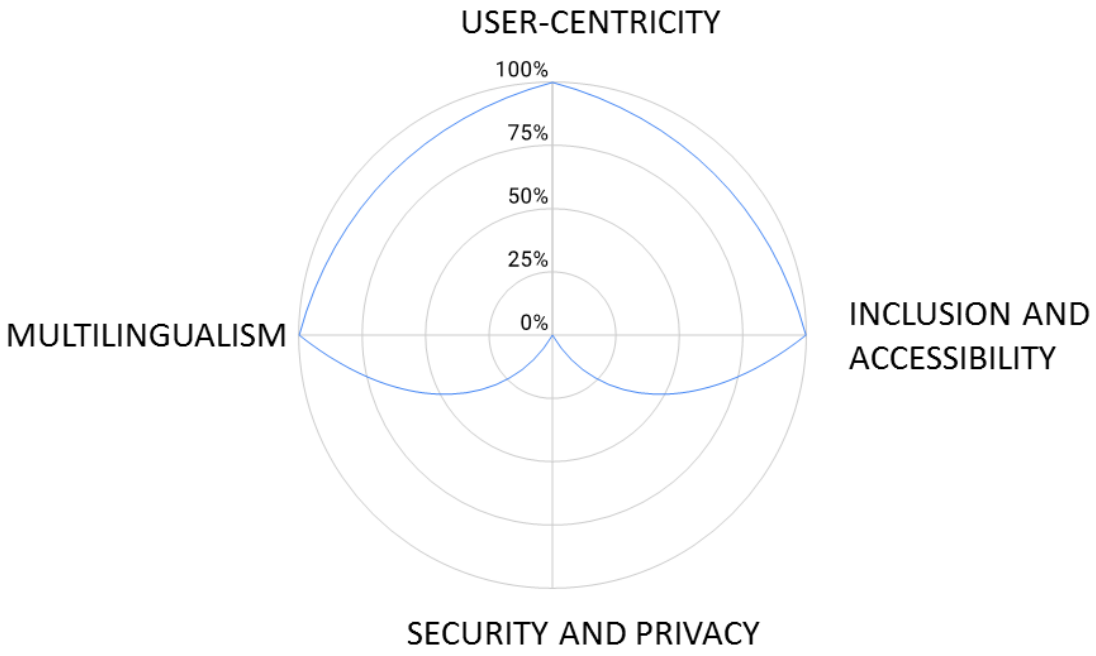
1.2.1 Orphanet standards

1.2.1.1 Results in graphs (in percentages)

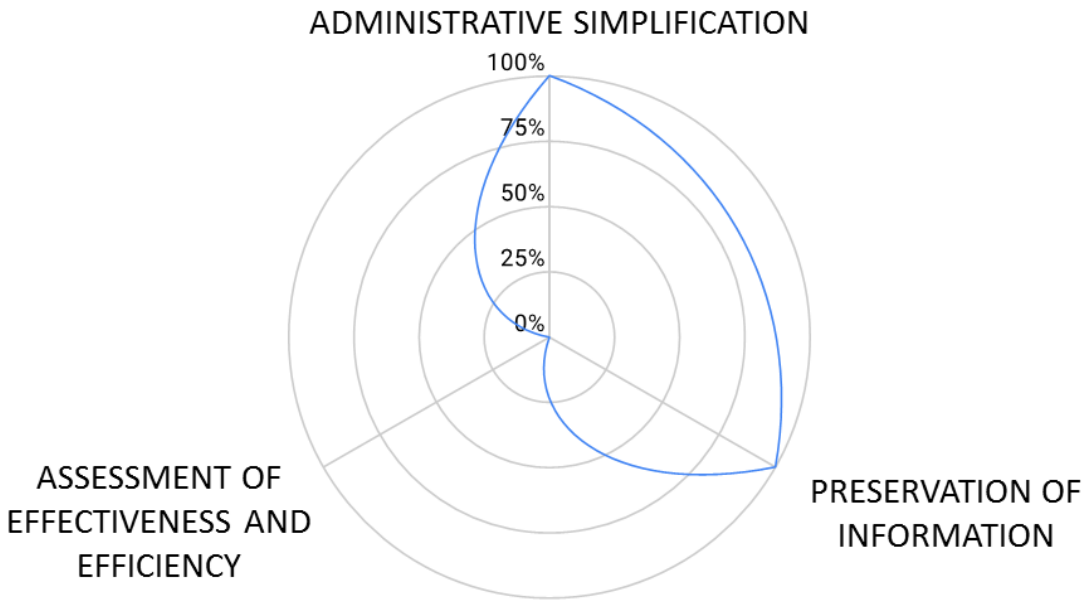
Core interoperability principles



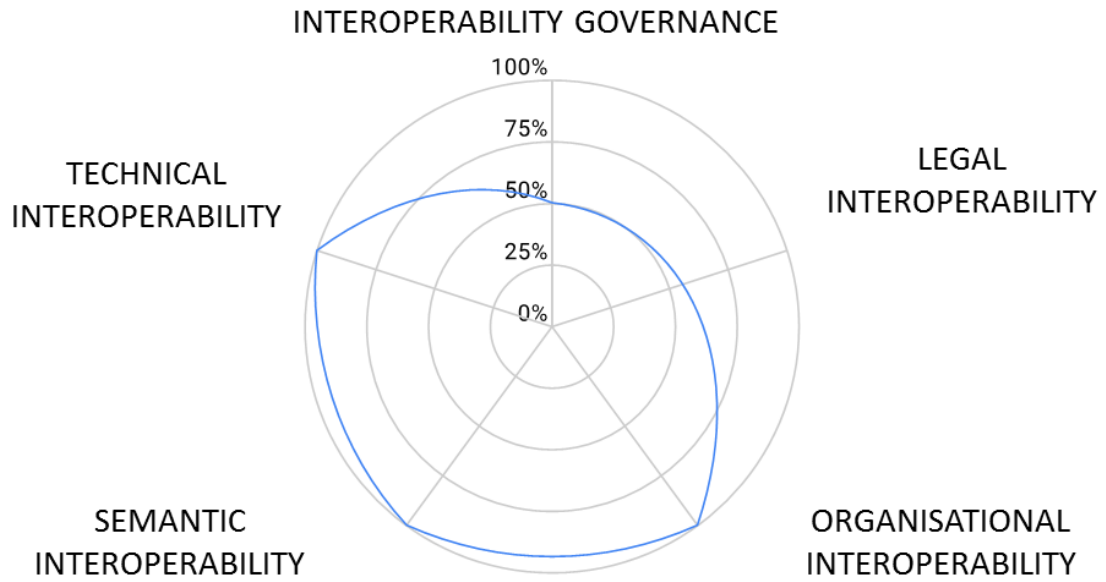
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.2.1.2 Interpretation according to compliance with EIF

The overall compliance of Orphanet standards with EIF is fairly high.

For the core interoperability principles dimension, there is full accomplishment for technological neutrality and data portability, whereas there is only partly accomplishment with respect to the openness dimension, although still high. This is mainly due to the fact that there is not sufficient market acceptance for its use in the development of products and services (A8). As far as transparency is concerned there is high compliance, since it has been judged that the criterion regarding protections of personal data managed by Public Administration (A13) is not relevant.

Regarding the principles related to generic user needs and expectations there is for three of the dimensions a high compliance with EIF, whereas the criteria concerning security and privacy (A25 and A26) were considered not relevant in this context.

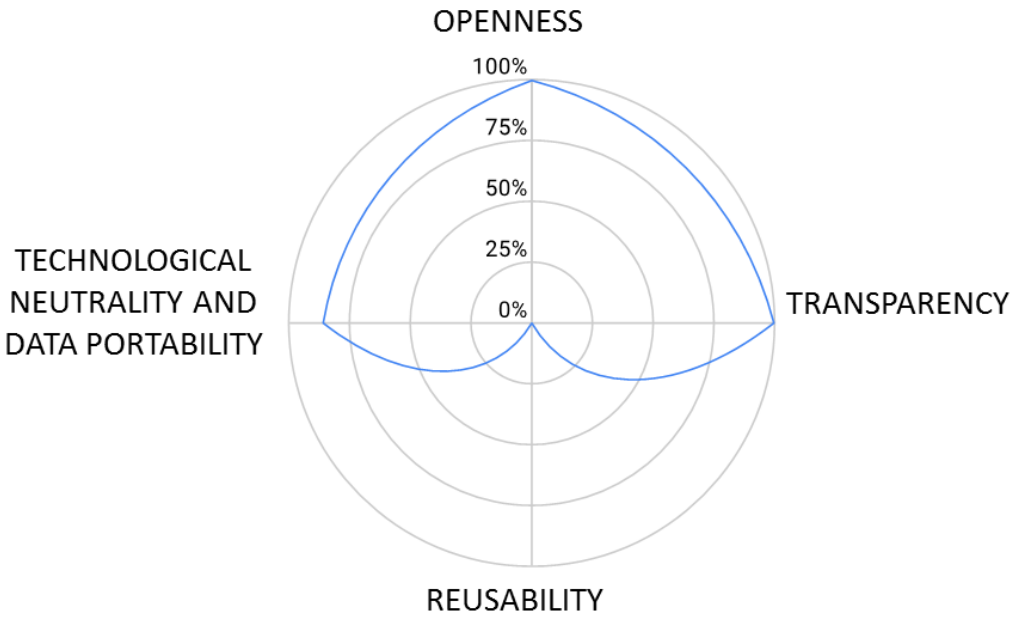
For the foundation principles for cooperation among public administrations there was found to be full compliance with EIF, except for the assessment of effectiveness and efficiency (A31 and A32) where no documentation has been provided.

In the interoperability layers, technical, semantic, and organisational interoperability fully comply with EIF, whereas interoperability governance only partly meets the criteria since there has been no use of the standard in an EU cross-border project or initiative after agreed identification and assessment (A36) and since the specification is not included in a repository/catalogue of standards at the national level (A37). It has not been judged relevant to consider whether the standard can be mapped to the European Interoperability Reference Architecture (A33). Finally, the legal operability criterion (A39) is not applicable.

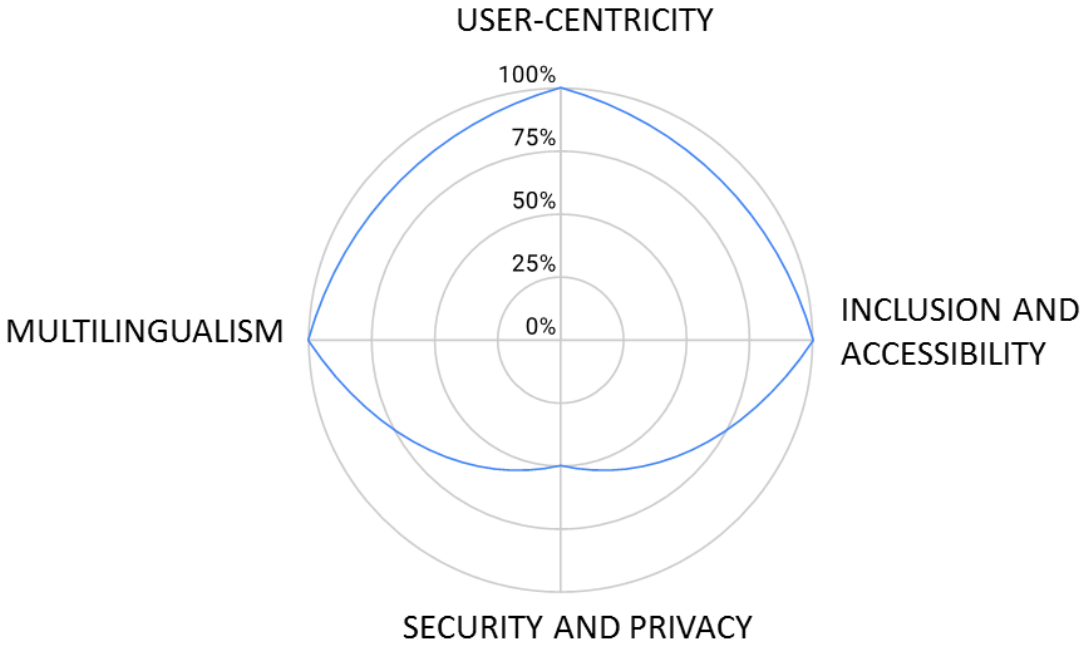
1.2.2 OMOP CDM

1.2.2.1 Results in graphs (in percentages)

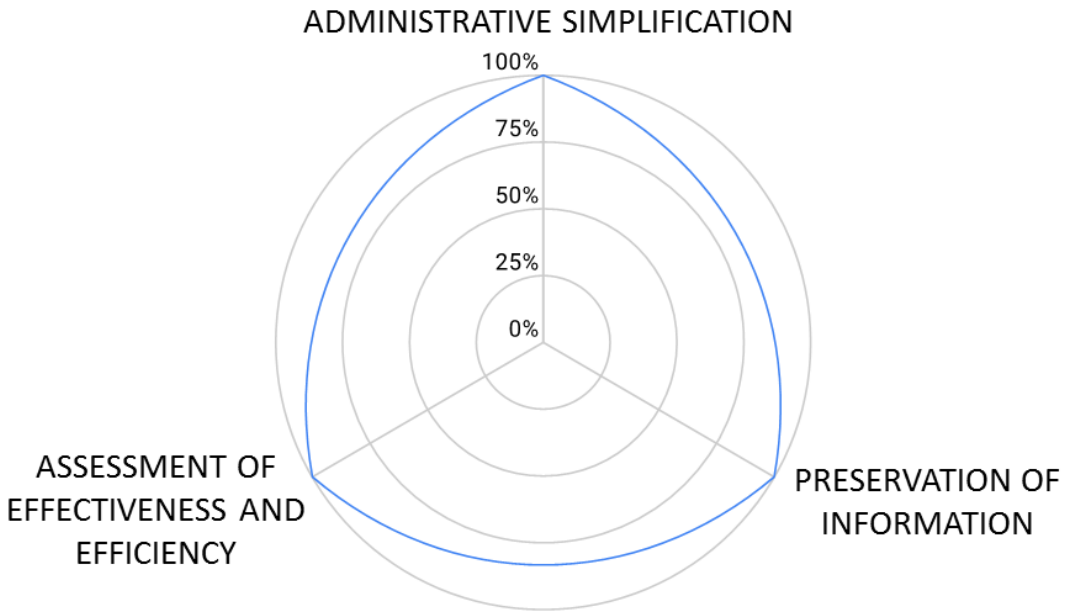
Core interoperability principles



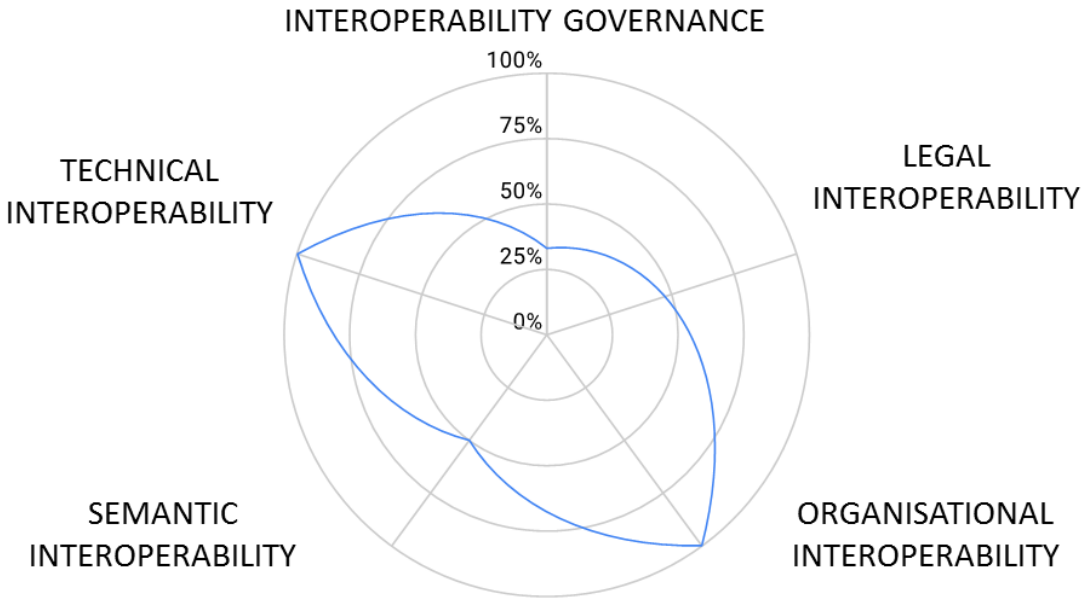
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.2.2.2 Interpretation according to compliance with EIF

The overall compliance of OMOP CDM with EIF is fairly high. The criteria within the foundation principles for cooperation among public administrations are fully compliant with EIF. On the other hand, the areas that scored lowest were the core interoperability principles and the interoperability layers.

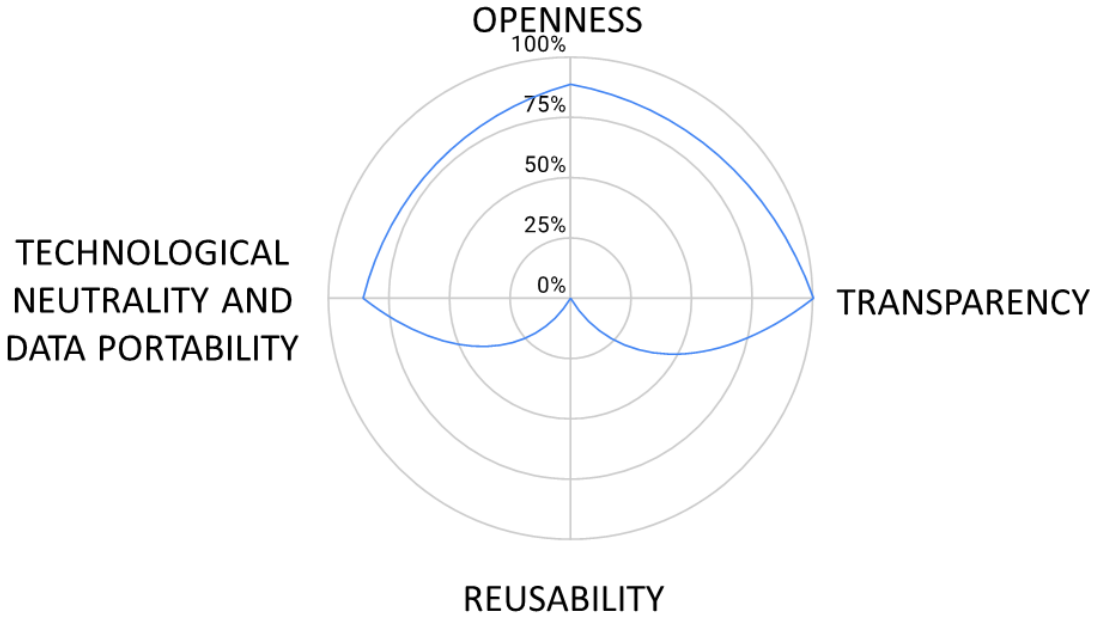
Reusability did not meet the criteria because it was assessed that the standard is not usable across business domains (A14, A15). The technological neutrality and data portability criterion was lowered because the standard does not allow partial implementations (A18). The interoperability governance criterion score was lower because it was assessed that mapping to EIRA had not been made (A33). The semantic interoperability criterion, on the other hand, was affected by the assessment that the standard does not foster the publication of data as linked open data (A43).

The security and privacy criterion was partly not relevant for this standard.

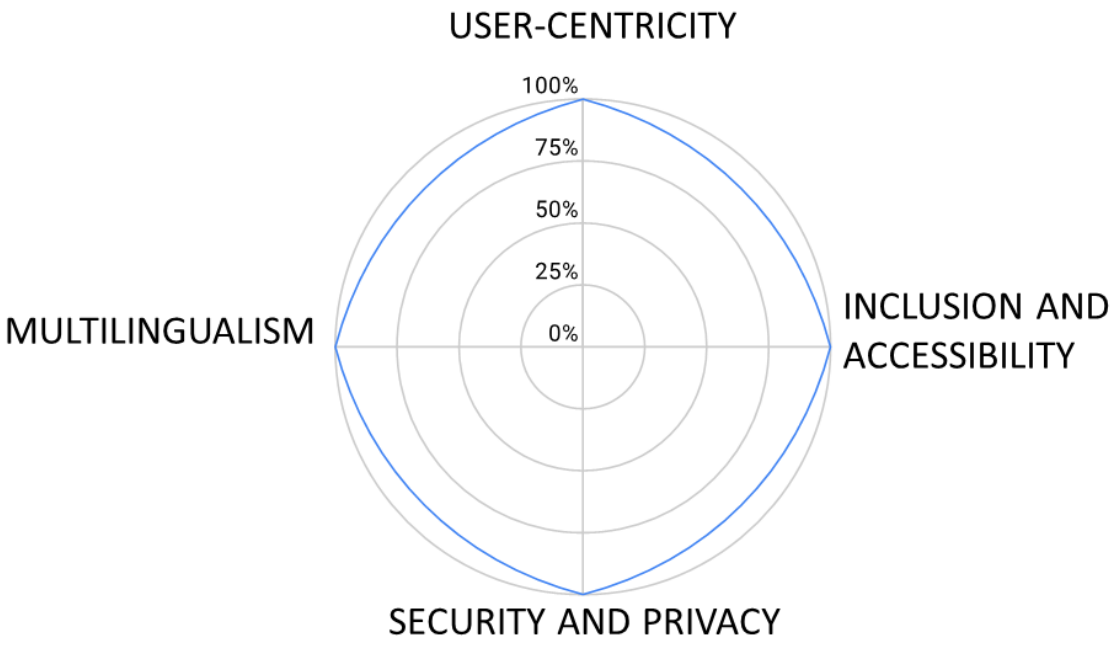
1.2.3 CDISC SDTM

1.2.3.1 Results in graphs (in percentages)

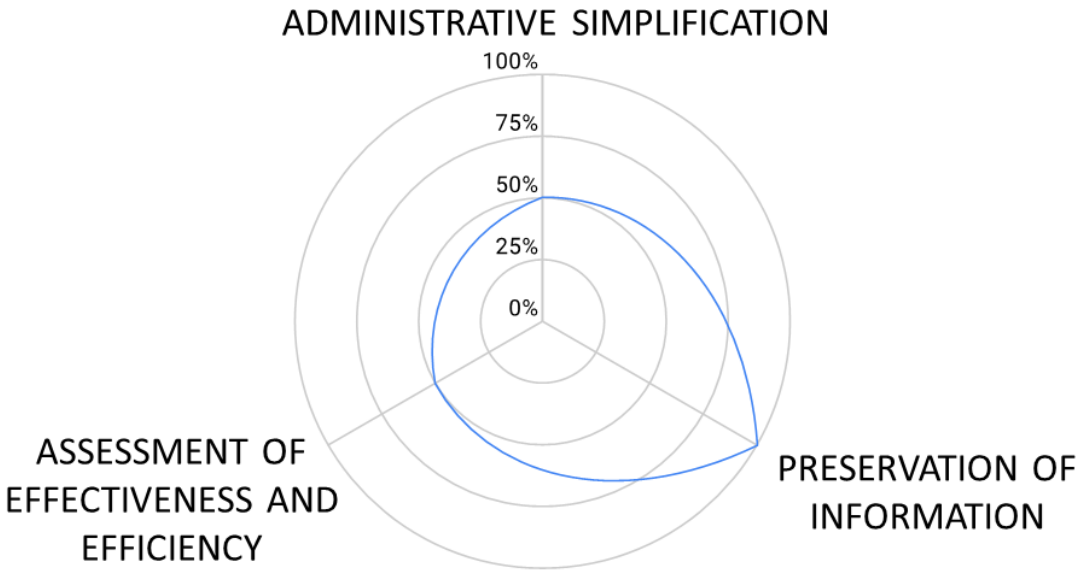
Core interoperability principles



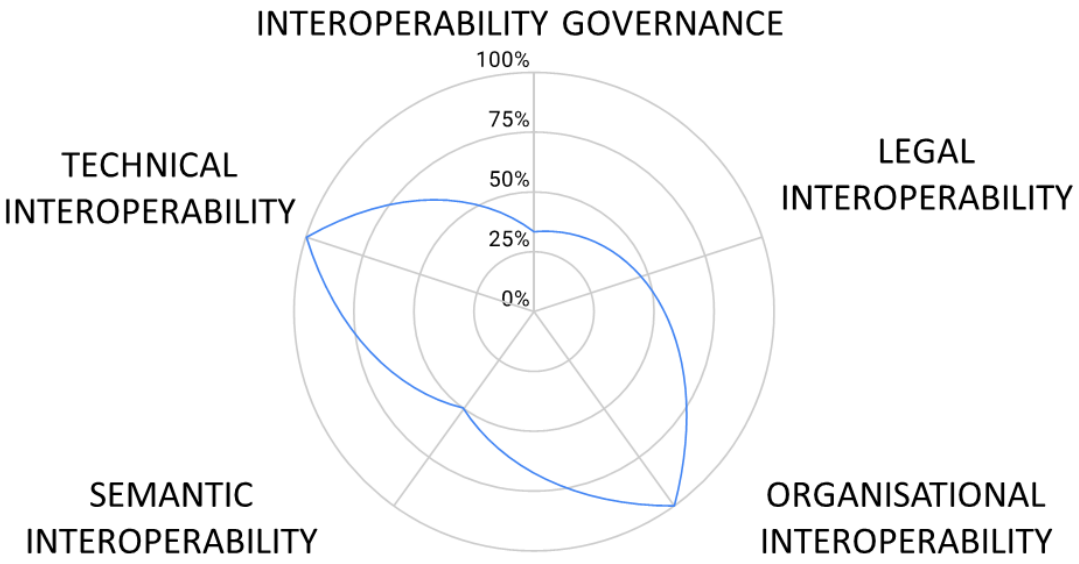
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.2.3.2 Interpretation according to compliance with EIF

The overall compliance of CDISC SDTM with EIF is quite good. The strengths are in the principles related to generic user needs and expectations including user-centricity.

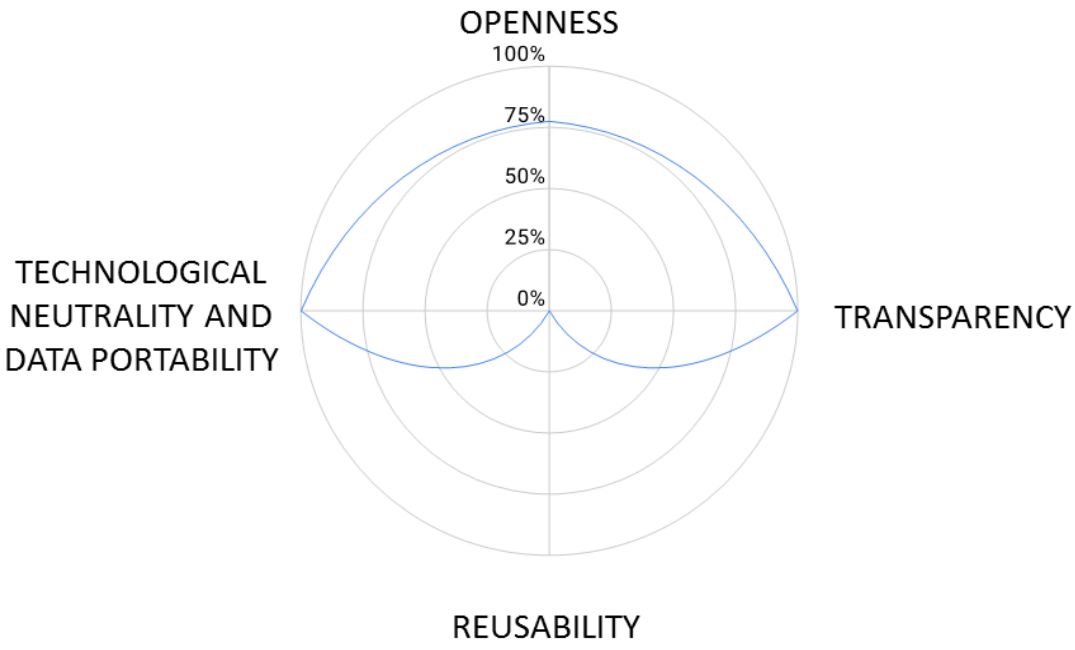
The weaknesses were found to be in reusability and also to some extent in openness, technological neutrality and data portability, administrative simplification, assessment of effectiveness and efficiency, interoperability governance and semantic interoperability.

The reusability criterion was not met because the standard is not usable beyond the business specific domain (A14, A15). The openness score was somewhat affected because the standard is not available for everyone to study without restrictions (A5). Semantic interoperability was affected because it was assessed that the standard doesn't foster the publication of data as linked open data (A43).

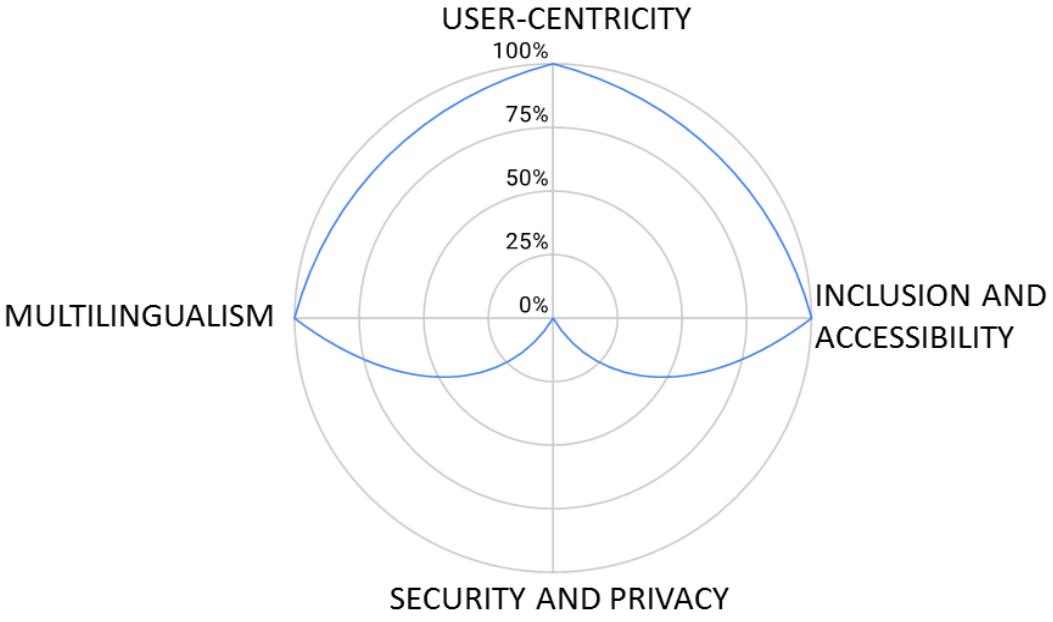
1.2.4 SNOMED CT

1.2.4.1 Results in graphs (in percentages)

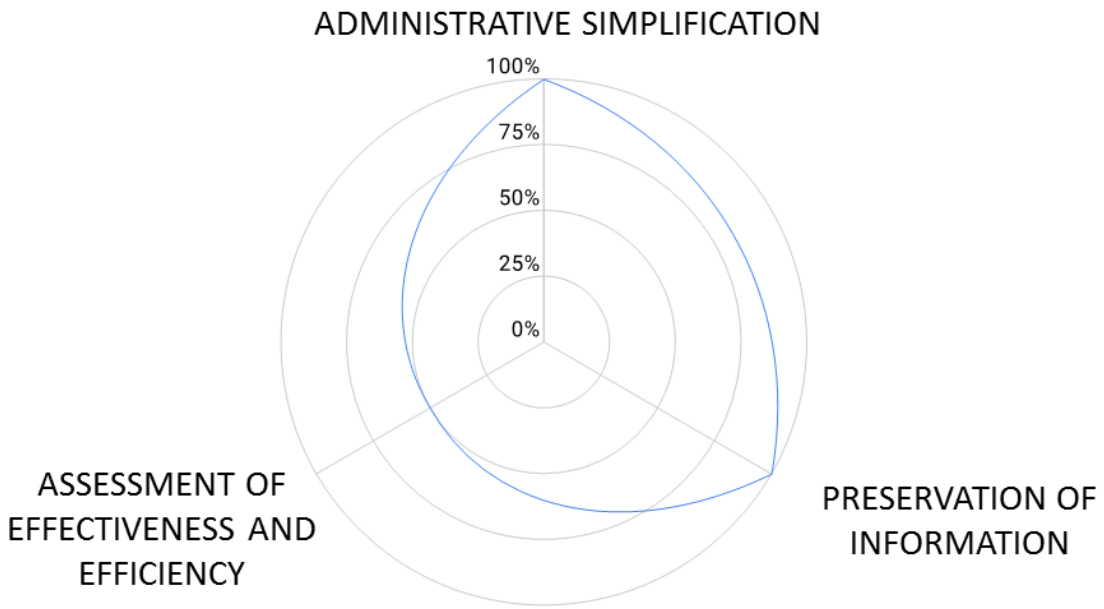
Core interoperability principles



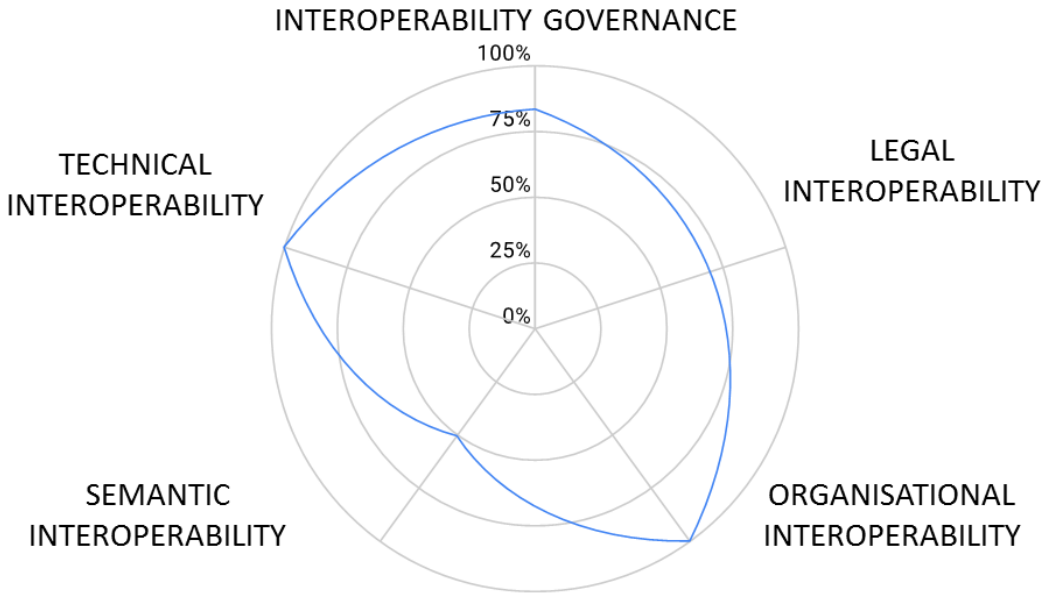
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.2.4.2 Interpretation according to compliance with EIF

The overall compliance of SNOMED CT with EIF is fairly high. All of the interoperability areas were found to be fairly even and there were no noteworthy differences in compliance between the principles.

However, the area that scored lowest was core interoperability principles. The weaknesses found were reusability and to an extent openness. SNOMED CT was not found to be reusable beyond a business-specific domain (A14, A15). Regarding the openness criterion, due to its licensing, this standard is unavailable to be studied with no restrictions (A5) and is not licenced as open source (A6), which lead to a lower evaluation score. Nonetheless, the current policy in the EU is promoting the implementation of SNOMED CT.

Within principles related to generic user needs and expectations, the security and privacy criterion (A25, A26) scored 0 because those questions were not relevant in the case of this standard. The other criteria scored highly.

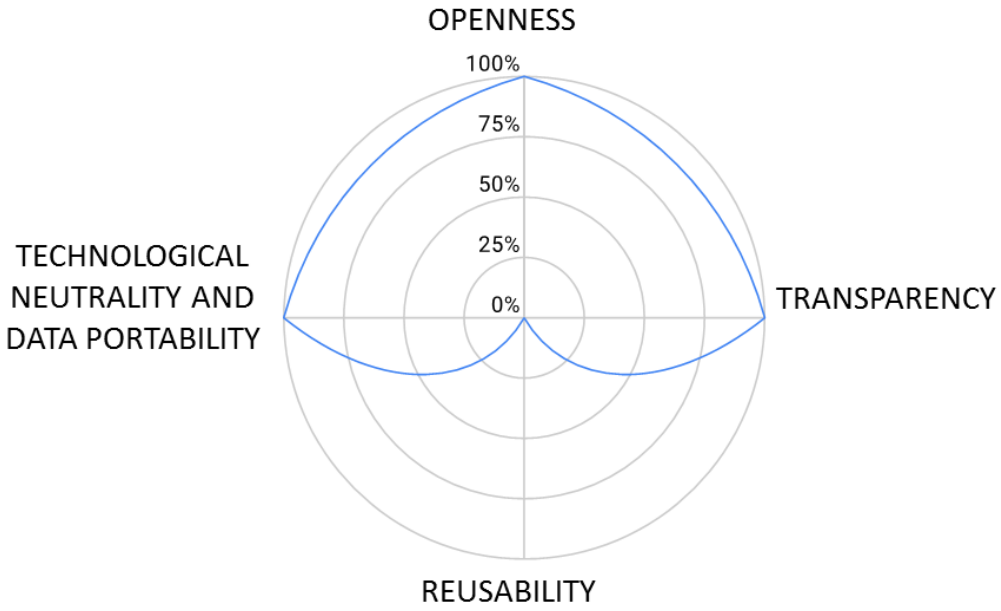
Within Foundation principles for cooperation among public administrations the weakness was found to be in assessments of effectiveness and efficiency criterion (A32) because studies assessing the efficiency could not be found.

In the interoperability layers group the criterion semantic interoperability lowered the score a bit because it was evaluated that SNOMED CT does not define a cross-sector reusable data model or encourage the sharing of the data in national platforms (A42).

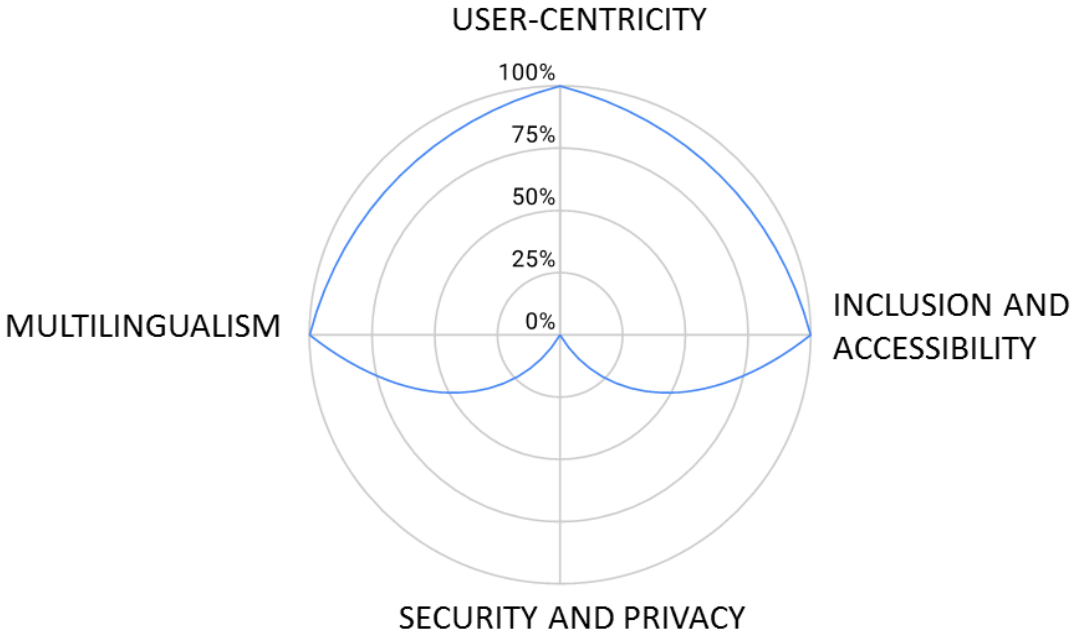
1.2.5 LOINC

1.2.5.1 Results in graphs (in percentages)

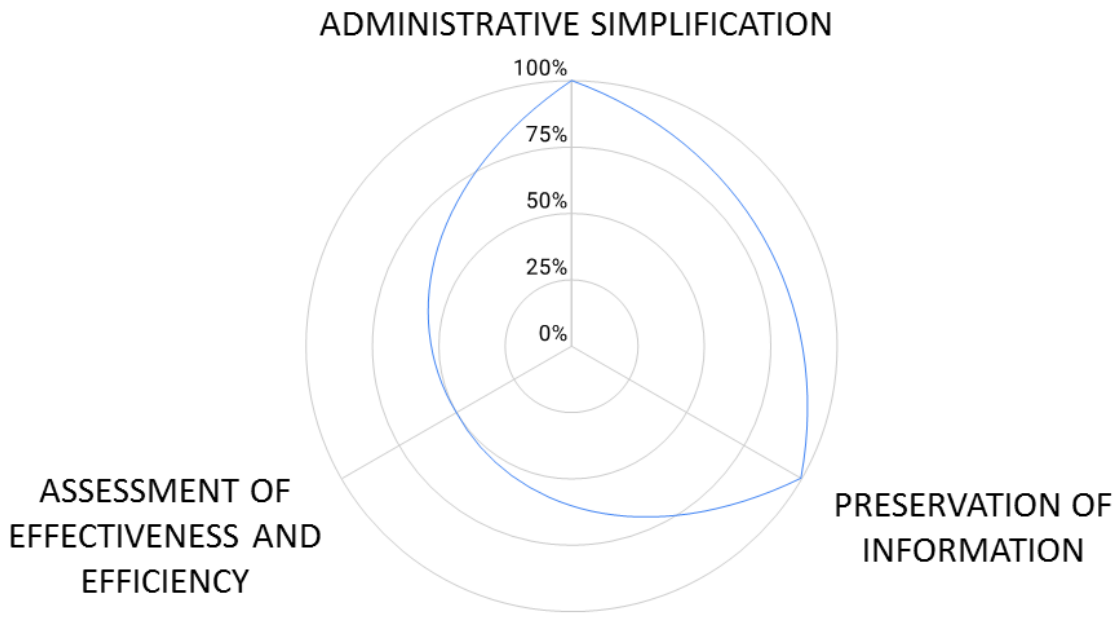
Core interoperability principles



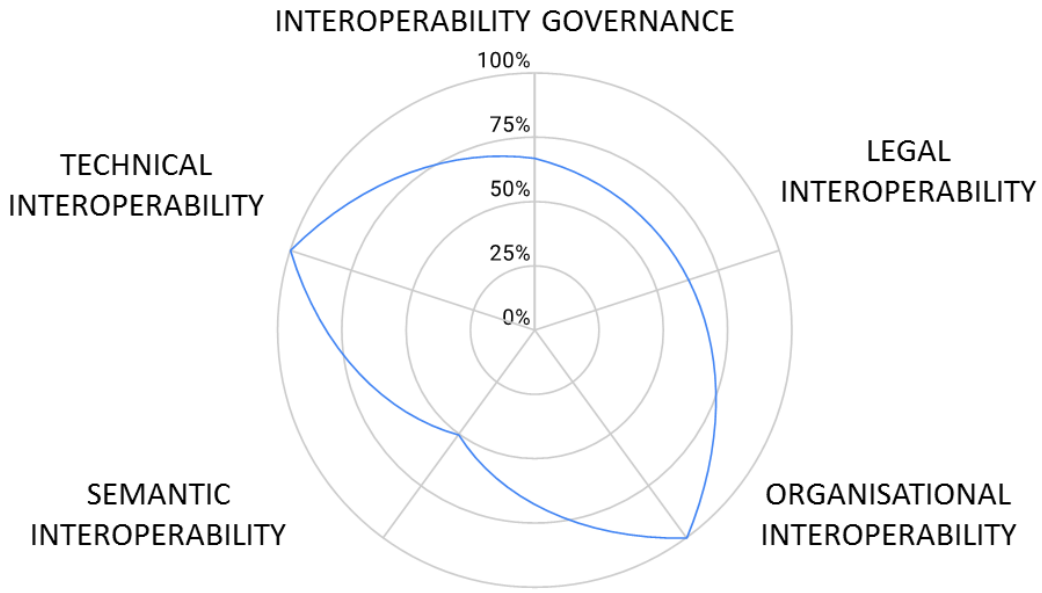
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.2.5.2 Interpretation according to compliance with EIF

LOINC was found to have a relatively high degree of compliance with EIF.

Within the core interoperability principles, openness, transparency, and technological neutrality and data portability fully fulfil the criteria and comply with EIF. Only reusability does not meet the criteria (A14 and A15).

The criteria within the principles related to generic user needs and expectations are fully met. The dimension concerning security and privacy is considered not to be relevant for this standard (A25 and A26).

Regarding the foundation principles for cooperation among public administrations, the criteria associated with administrative simplification and preservation of information are completely fulfilled. However, no documentation of assessment of efficiency (A32) has been found, whereas assessment of effectiveness (A31) has.

In the interoperability layers, technical interoperability and organisational interoperability fully meet the criteria, whereas semantic interoperability only partly does since the specification does not encourage the creation of communities along with the sharing of their data (A43). For the interoperability governance dimension, the evaluation is lower since it is judged that the conformance of the specification's implementations (A34) cannot be assessed. It should also for this dimension be remarked that it is not considered relevant whether the standard can be mapped to the European Interoperability Reference Architecture (A33). Finally, the criterion concerning legal interoperability is not applicable (A39).

1.3 Standards for interoperable communication

The standards in this group are DICOM, HL7 FHIR, IDMP-SPOR and ISO 8000 110. The average scores for each are shown in table 3.

Table 3. Average scores for messaging between nodes standards

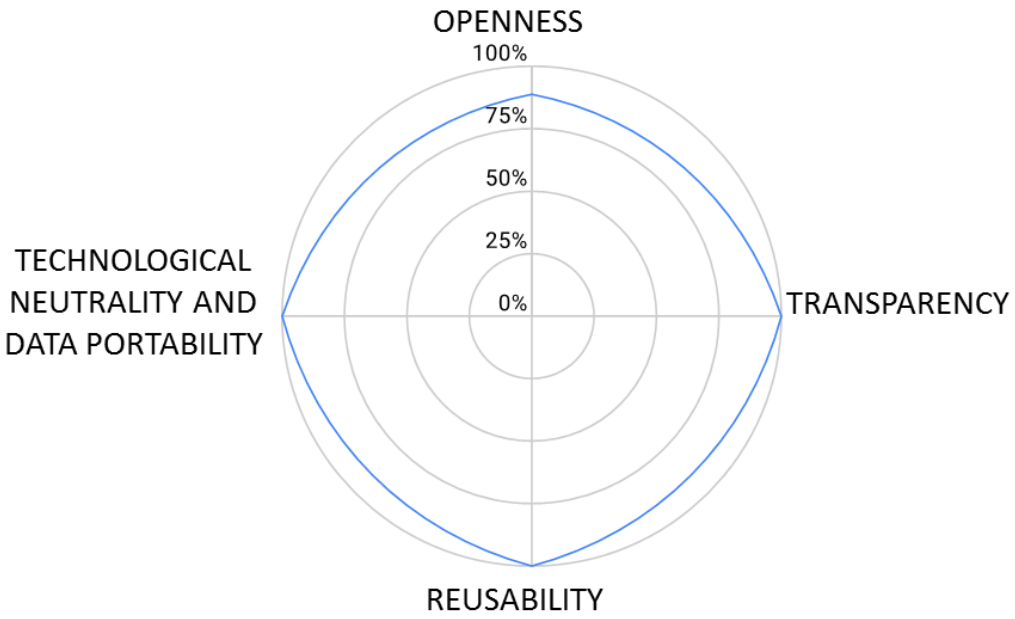
	CRITERION	DICOM	HL7-FHIR	IDMP	ISO 8000-110
CORE INTEROPERABILITY PRINCIPLES		100%	100%	44%	97%
	OPENNESS	100%	100%	78%	89%
	TRANSPARENCY	100%	100%	25%	100%
	REUSABILITY	100%	100%	0%	100%
	TECHNOLOGICAL NEUTRALITY AND DATA PORTABILITY	100%	100%	71%	100%
USERS' NEEDS AND EXPECTATIONS		100%	100%	50%	100%
	USER-CENTRICITY	100%	100%	100%	100%

	INCLUSION AND ACCESSIBILITY	100%	100%	100%	100%
	SECURITY AND PRIVACY	100%	100%	0%	100%
	MULTILINGUALISM	100%	100%	0%	100%
PRINCIPLES FOR COOPERATION BETWEEN INSTITUTIONS		100%	100%	67%	67%
	ADMINISTRATIVE SIMPLIFICATION	100%	100%	100%	100%
	PRESERVATION OF INFORMATION	100%	100%	100%	100%
	ASSESSMENT OF EFFECTIVENESS AND EFFICIENCY	100%	100%	0%	0%
INTEROPERABILITY LAYERS		96%	96%	67%	96%
	INTEROPERABILITY GOVERNANCE	83%	83%	67%	83%
	LEGAL INTEROPERABILITY	N/A	N/A	N/A	N/A
	ORGANISATIONAL INTEROPERABILITY	100%	100%	50%	100%
	SEMANTIC INTEROPERABILITY	100%	100%	50%	100%
	TECHNICAL INTEROPERABILITY*	100%	100%	100%	100%
	OVERALL SCORE	99%	99%	57%	90%

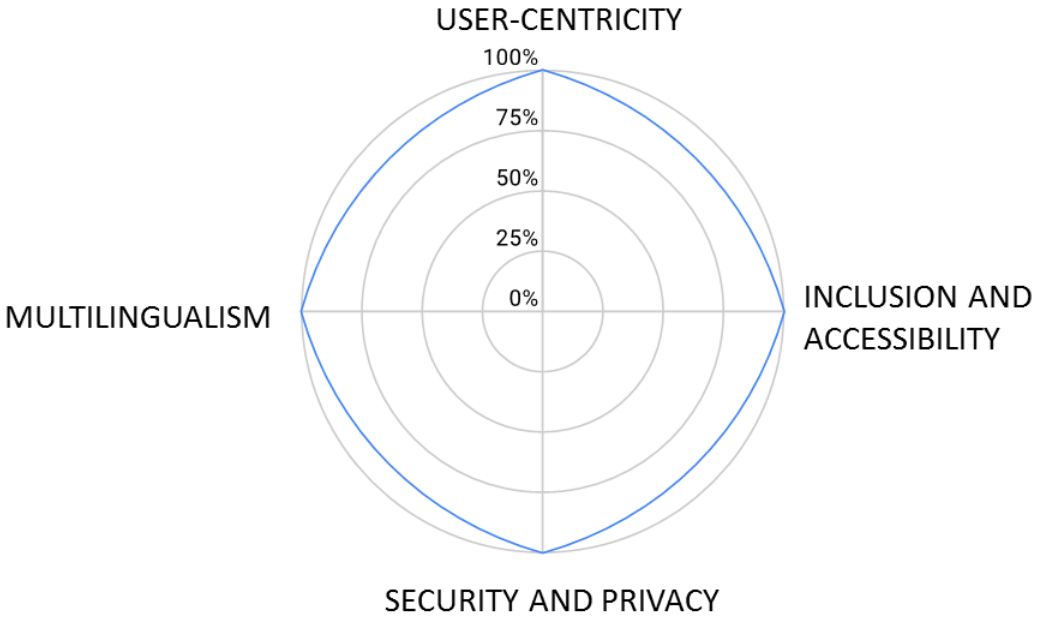
1.3.1 ISO 8000 110

1.3.1.1 Results in graphs (in percentages)

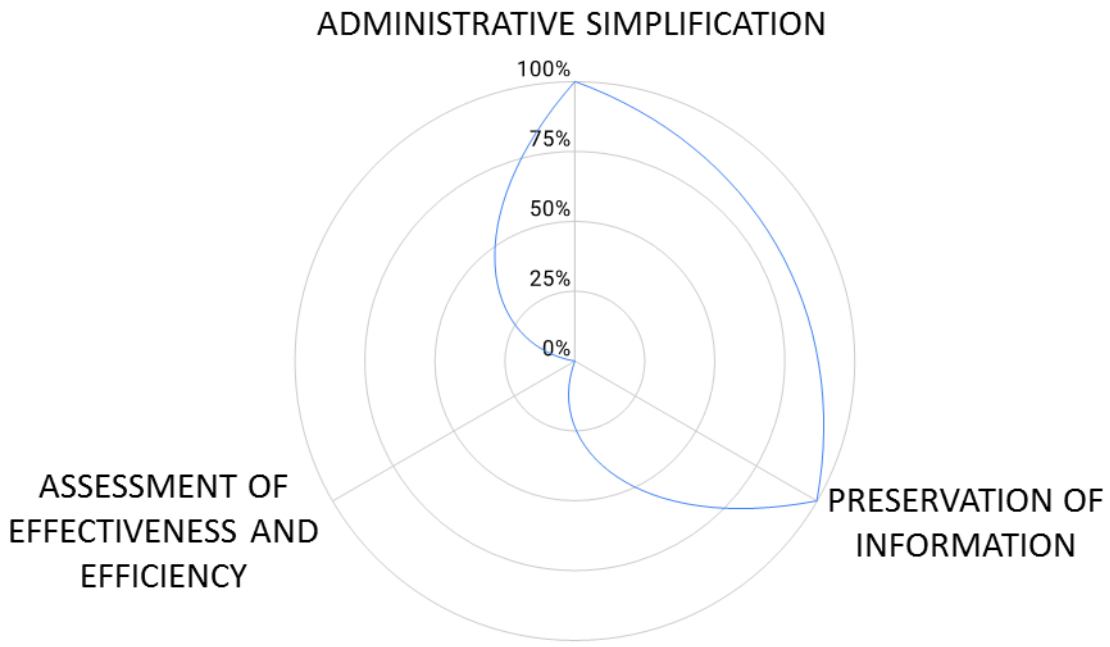
Core interoperability principles



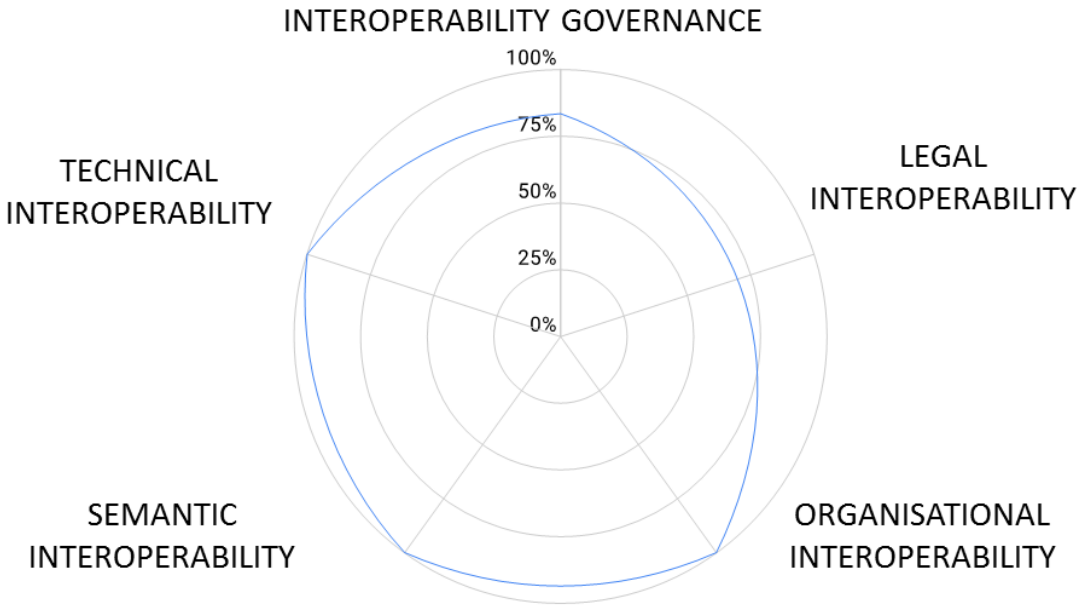
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.3.1.2 Interpretation according to compliance with EIF

ISO 8000 110 was found to have a high degree of compliance with EIF.

Within the core interoperability principles, only openness does not fully meet the criteria - since the specification is not available for everyone to study (A5). All other core interoperability criteria (transparency, technological neutrality and data portability, as well as reusability) comply with EIF.

The criteria within the principles related to generic user needs and expectations are all fully met.

Regarding the foundation principles for cooperation among public administrations, the criteria associated with administrative simplification and preservation of information are completely fulfilled. However, no documentation of assessment of effectiveness and efficiency has been found.

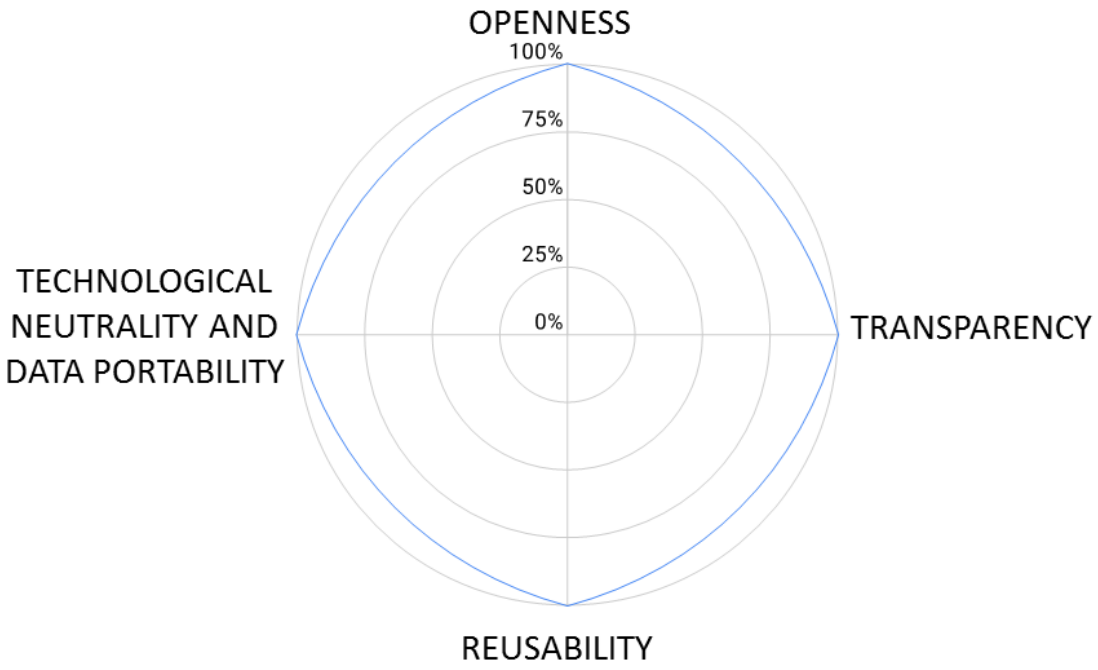
In the interoperability layers, technical interoperability, semantic interoperability, organisational interoperability, and interoperability governance comply with EIF. With respect to the latter, it should be remarked that it has not been considered relevant whether the standard can be mapped to the European Interoperability Reference Architecture (A33). Nevertheless, this criterion is still included in the radar graph, resulting in a value below 100%

for interoperability governance, although it is fully compliant with respect to the criteria taken into consideration. Additionally, the criterion concerning legal interoperability is not applicable (A39).

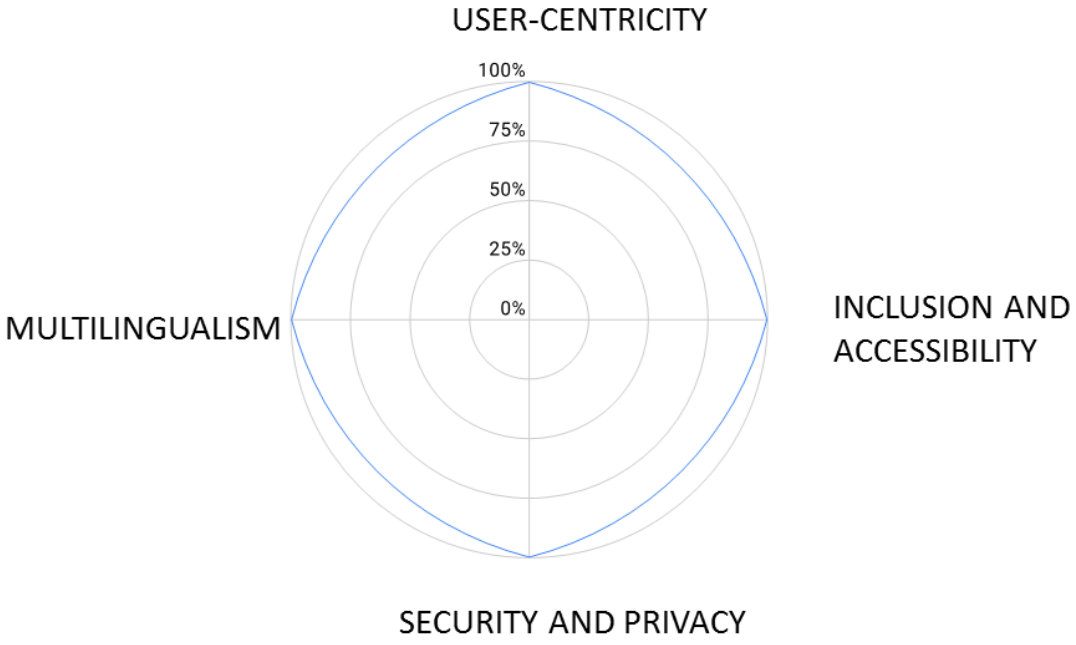
1.3.2 HL7 FHIR

1.3.2.1 Results in graphs (in percentages)

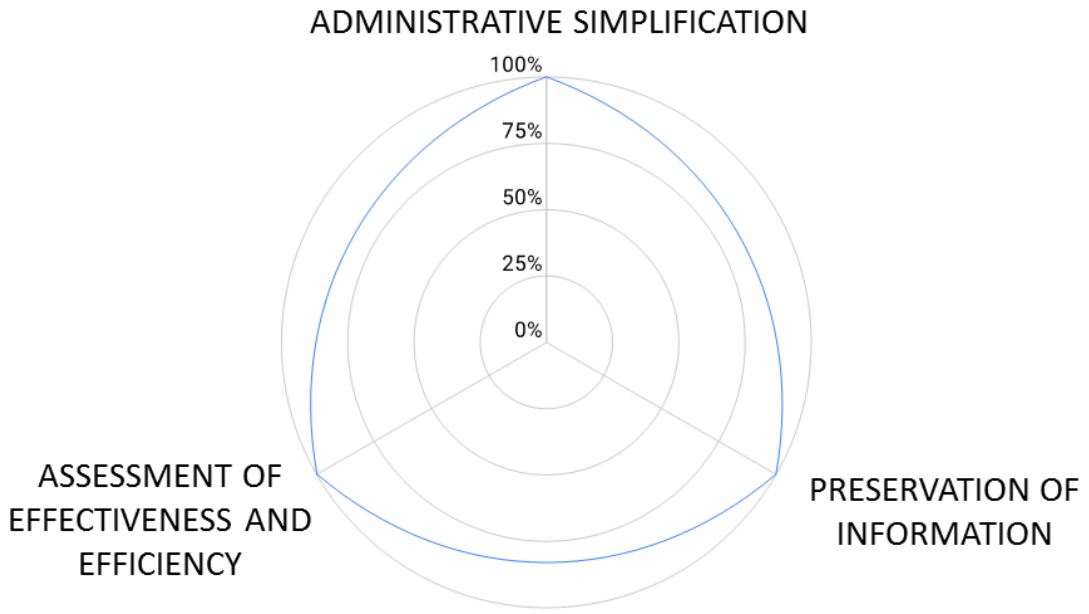
Core interoperability principles



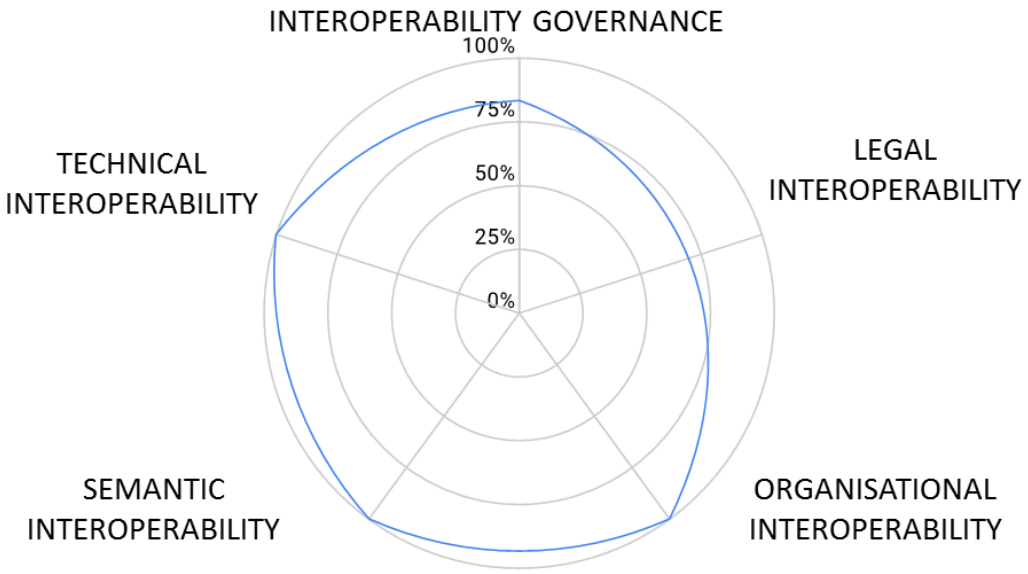
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.3.2.2 Interpretation according to compliance with EIF

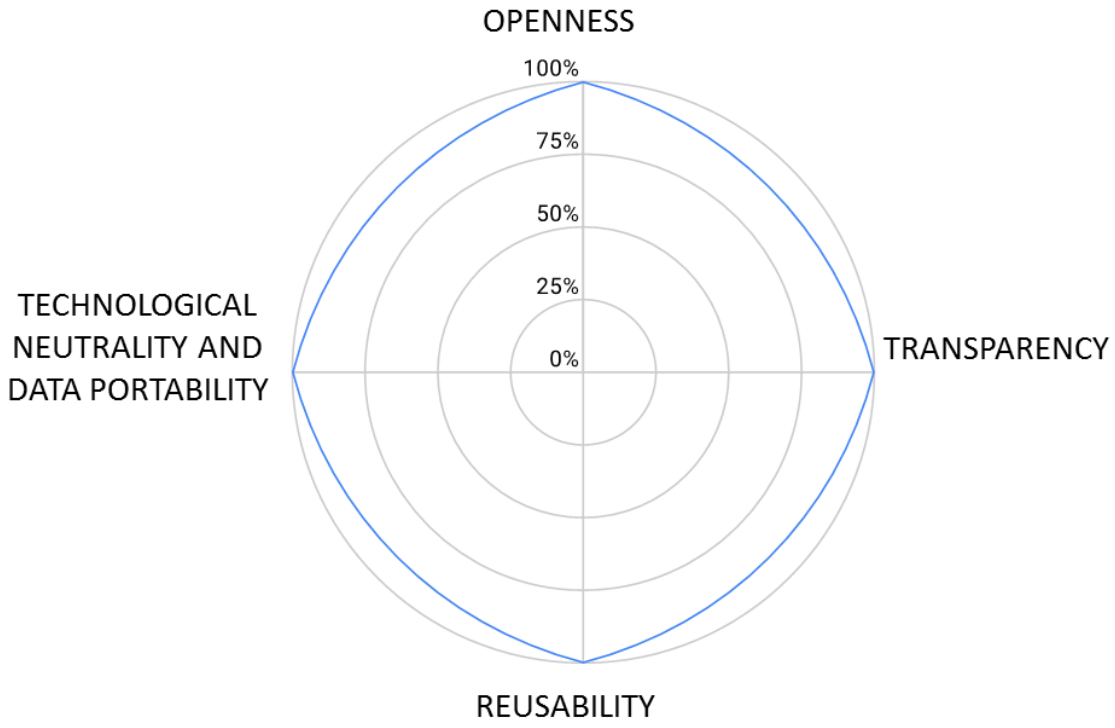
HL7 FHIR has full compliance with EIF with respect to all dimensions considered within the core interoperability principles, principles related to generic user needs and expectations, the foundation principles for cooperation among public administrations, and the interoperability layers.

However, the interoperability layers need some commenting. It should be remarked that it has not been considered relevant whether the standard can be mapped to the European Interoperability Reference Architecture (A33). Nevertheless, this criterion is still included in the radar graph resulting in a value below 100% for interoperability governance, although it is fully compliant with respect to the criteria taken into consideration. Additionally, the criterion concerning legal interoperability is not applicable (A39).

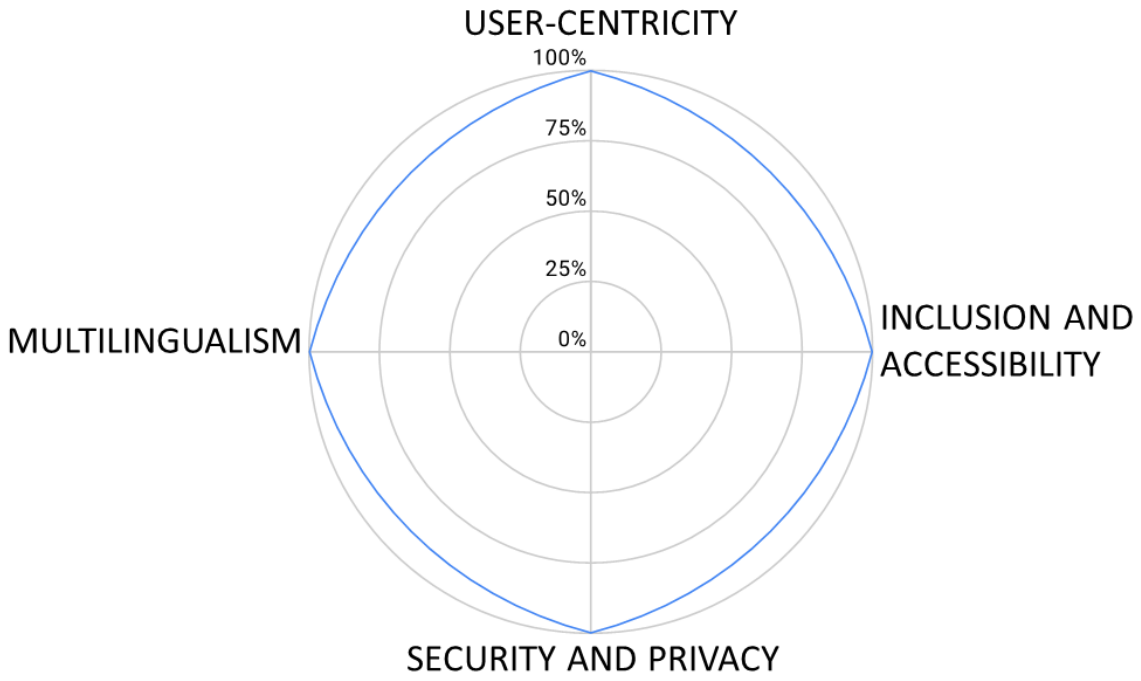
1.3.3 DICOM

1.3.3.1 Results in graphs (in percentages)

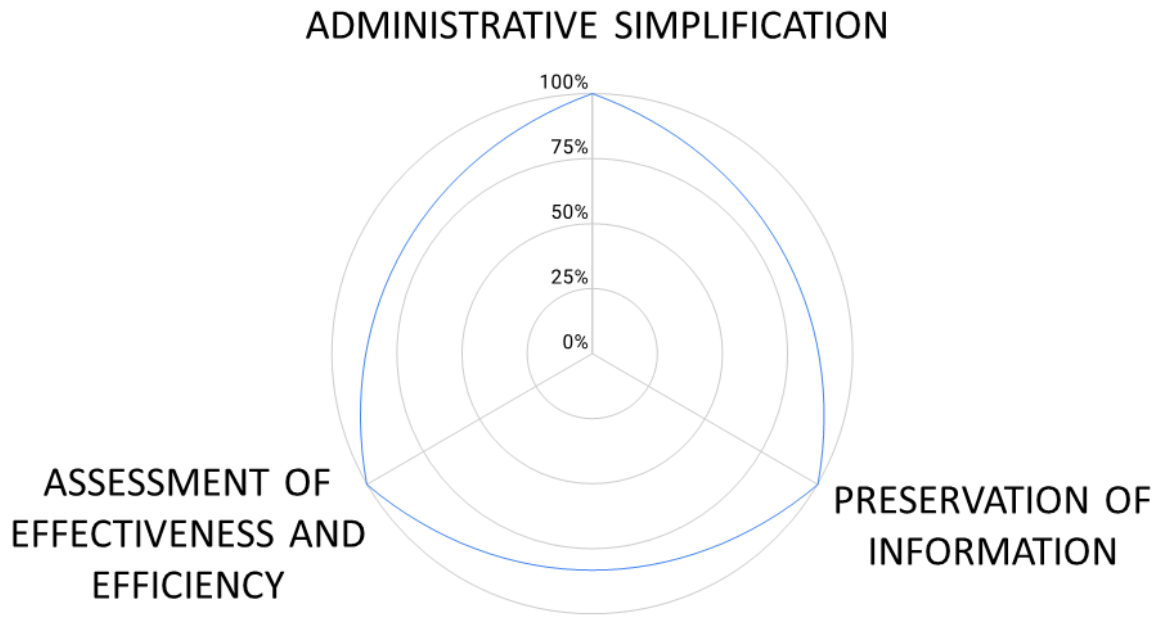
Core interoperability principles



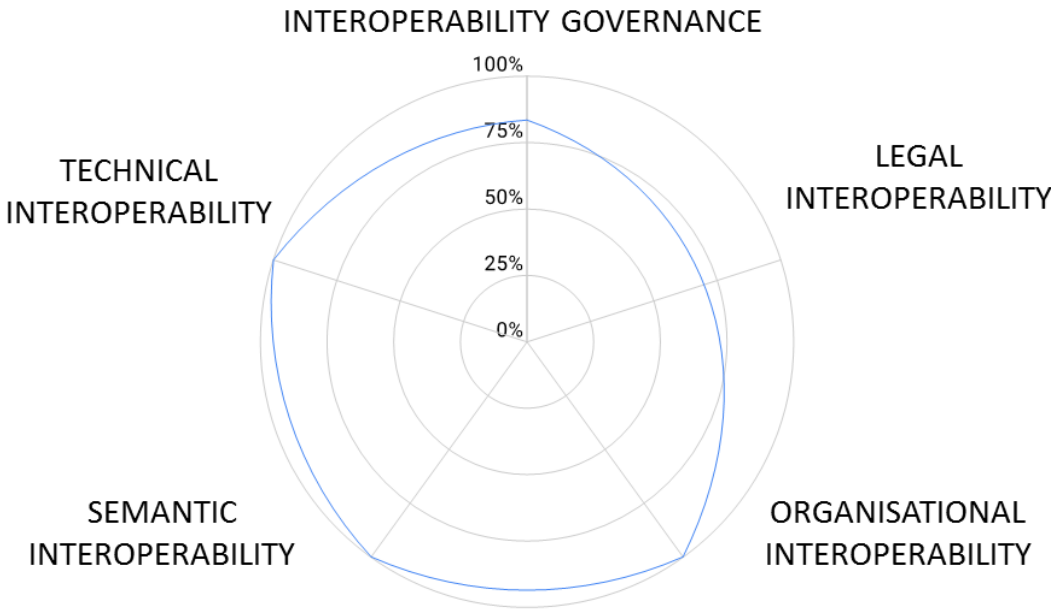
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.3.3.2 Interpretation according to compliance with EIF

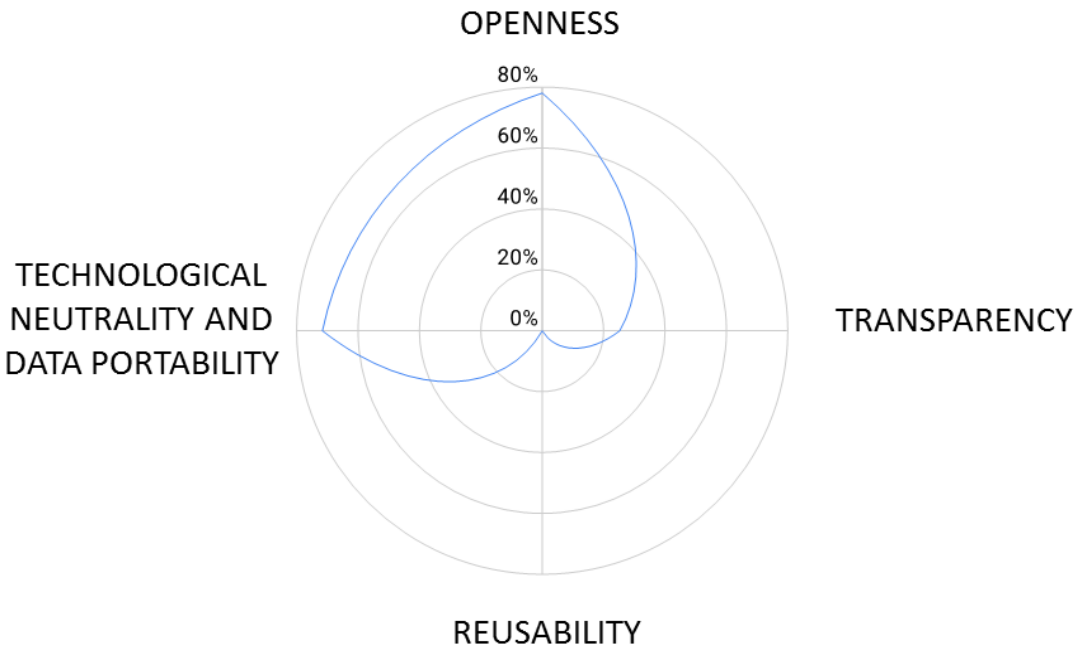
DICOM has full compliance with EIF with respect to all dimensions considered within the core interoperability principles, principles related to generic user needs and expectations, the foundation principles for cooperation among public administrations, and the interoperability layers.

However, the interoperability layers need some commenting. It should be remarked that it has not been considered relevant whether the standard can be mapped to the European Interoperability Reference Architecture (A33). Nevertheless, this criterion is still included in the radar graph resulting in a value below 100% for interoperability governance, although it is fully compliant with respect to the criteria taken into consideration. Additionally, the criterion concerning legal interoperability is not applicable (A39).

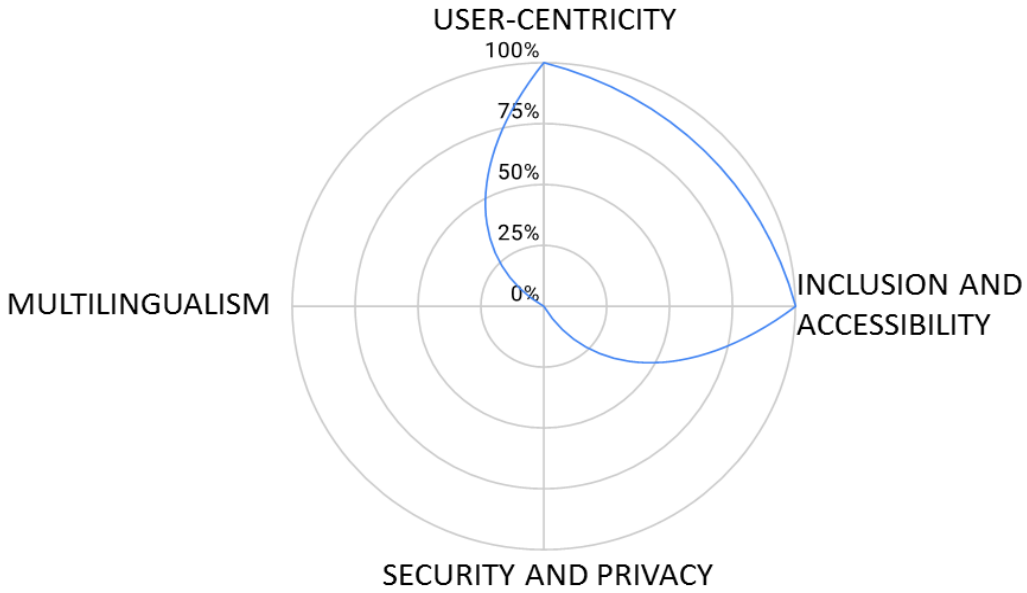
1.3.4 SPOR

1.3.4.1 Results in graphs (in percentages)

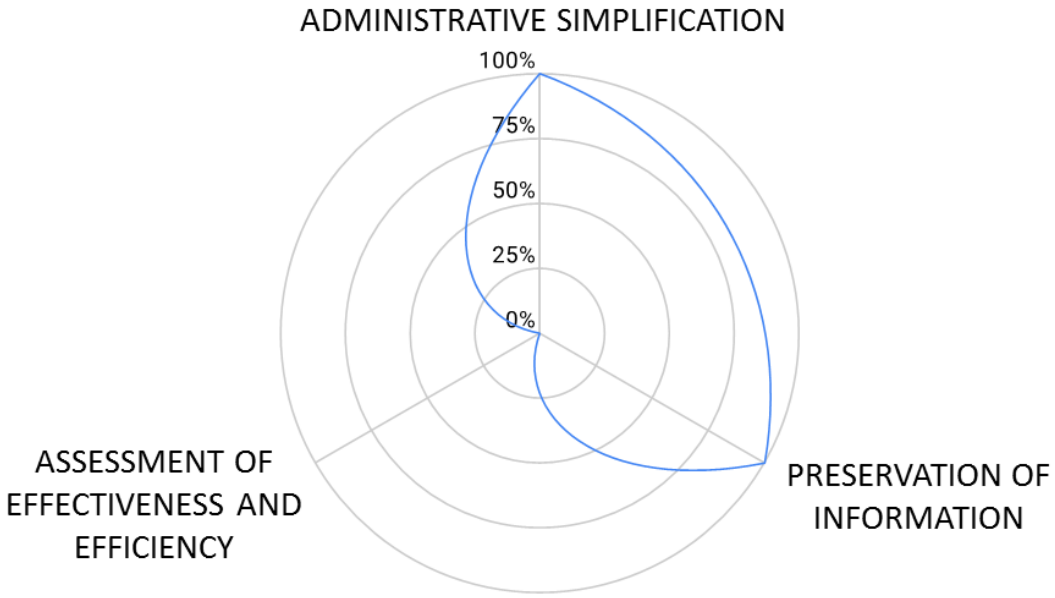
Core interoperability principles



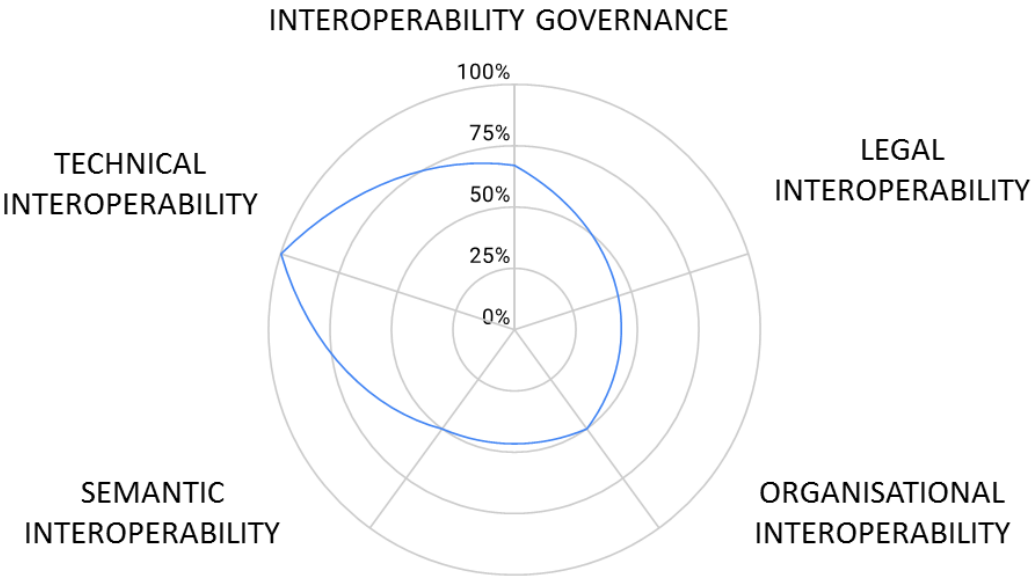
Principles related to generic user needs and expectations



Foundation principles for cooperation among public administrations



Interoperability layers



1.3.4.2 Interpretation according to compliance with EIF

The overall compliance of SPOR with EIF is not so high.

The criteria regarding the core interoperability principles are not fully met. This can primarily be attributed to a public review not being part of the decision-making process (A4) and the standard not being licensed on a royalty-free basis (A6), reflected in the openness dimension, the specification not ensuring the protection of personal data (A13), reflected in the transparency dimension, no existence of reusability features (A14 and A15), as well as no possibility of neither partial implementation (A18) nor customisation (A19), reflected in the technological neutrality and data portability dimension.

The principles related to generic user needs and expectations score high for user-centricity, as well as inclusion and accessibility, but low for multilingualism (A27). Security and privacy criteria (A25, A26) are not relevant for this standard.

In general, the foundation principles for cooperation among public administrations have a positive evaluation and meet the criteria - with the exception of assessment of effectiveness and efficiency (A31 and A32) since no supporting documentation has been found.

The interoperability layer is assessed to comply with EIF on an intermediate level. It scores highest for technical interoperability (full compliance), but is only intermediate for organisational interoperability and semantic operability due to not facilitating the modelling of business processes (A40) and not encouraging the creation of communities along with the sharing of data and results on national platforms (A42). Within interoperability governance, it can be remarked that it has not been considered relevant whether the specification can be mapped to the European Interoperability Reference Architecture (A33). Moreover, it has been assessed that the conformance of the specification's implementations cannot be assessed (A34). Finally, legal interoperability (A39) is not applicable for this standard.

Section 2. Implementation - country survey results

Partners in WP6 were asked about their actual experience within their countries, how widely standards were adopted, whether they had a role in the National Framework (if any) and what would be the main barriers for implementation.

In the following tables, there is a distillation of the responses:

Experience within the countries

In annex III the responses gathered from the different countries participating in the survey may be found.

Experience on standards for data discoverability

None of the countries has experience on all the standards evaluated in this report. On the other hand, none of the standards is being used in all the surveyed countries.

The experience in those cross-domain standards, DCAT-AP, INSPIRE and FAIRSHARING, has been proven very limited in this sample of countries. DCAT-AP is mentioned in France, Finland (with an *ad hoc* extension) and Norway; INSPIRE in the Czech Republic, Finland, and Norway, and FairSharing is named in the Czech Republic.

Out of that domain specific, BBMRI-MIABIS and ECRIN are the only ones named in a substantive number of countries - the former in Austria, the Czech Republic, France, Finland, Hungary, the Netherlands and Norway; and, the latter in the Czech Republic, France, Hungary and Norway.

Are they part of the National Framework or widely adopted?

None of these standards has been adopted as part of a National Framework. Only the Norwegian Directorate of eHealth has developed a metadata specification more or less based on DCAT-AP properties and all health data sources in Norway have to share their metadata according to this specification.

It is also worth mentioning that even though INSPIRE requires member states the publication of geo-data following the Standard, only the Norwegian Mapping Authority is mentioned as having the standard in a National Framework.

Although not necessarily under the consideration of a National Framework, those countries that act as members of research infrastructures, such as BBMRI-ERIC or ECRIN, have adopted MIABIS and CRMDR as part of their commitments.

Experience on standards enabling semantic interoperability

In the case of standards that could help build a layer of semantic interoperability, all the interviewees have declared having experience in SNOMED CT. Except Denmark, Norway, and Sweden, all the countries declared having experience in LOINC; and all, except Denmark and Ireland, declared having experience in Orphanet standards. OMOP-CDM is mentioned in all the countries except the Czech Republic, Denmark, Hungary, Ireland, Sweden, and Portugal. Interestingly the use of OMOP CDM is currently being discussed as a standard for a national data repository in Austria, and the German EHRs will be made available in the OMOP CDM for secondary use (still at an early stage). Fewer countries declared experience in CDISC - Austria, France, Germany and Sweden.

Are they part of the National Framework or widely adopted?

As per the responses, SNOMED CT, although partially, has been implemented in Austria, Denmark, France, Finland, Germany, Ireland, Portugal, Spain, and Sweden; LOINC (Austria, France, Germany, Hungary, The Netherlands, Portugal, Spain, and partly, in Finland and Ireland); Orphanet standards (France, Germany, Portugal, partly in Finland, and mapped out to SNOMED CT in Spain); and OMOP-CDM (in France, and the Norwegian Directorate of eHealth is fostering a task in the Nordic Commons Project). CDISC is not mentioned as being even partially used.

Experience on standards enabling interoperable communication

The analysis of the responses on the experience of standards that enable semantic and syntactic communication across nodes showed that 1) all the countries have experience in HL7 FHIR. Interestingly, it is expected to have a wide adoption in the upcoming years in Austria, the Czech Republic and Denmark. Except Germany, all the surveyed declared experience in DICOM. Experience in IDMP as a standard for communication is declared in half of the countries - France, Finland, Germany, Hungary and Sweden do not declare any experience. ISO 8000-110 is believed to be used in Denmark, and ORPHANET-INSERM (France).

Are they part of the National Framework or widely adopted?

DICOM is deemed part of the National Framework or widely adopted in all the countries except Germany. HL7 FHIR has been widely adopted in Denmark, France, Finland, Hungary, Portugal, the Czech Republic, and the Netherlands. In the case of Spain, it is the predecessor HL7 CDA (also in Finland), although used for primary purposes. In the case of Norway there is an increasing use. Finally, ISO 8000-110 is just highly adopted in ORPHANET INSERM.

Barriers and challenges for implementation

In this section, just those comments on barriers and challenges are included. Text is referred *verbatim*

Barriers and challenges on standards for data discoverability

Standard	Comments (country)
Beacon	N/A
BBMRI-MIABIS	
Bio-image archive	
CESSDA CMM	
DCAT-AP	*The HDH metadata catalogue has three levels of database description: base level, table level, variable level. DCAT does not have a variable level equivalent. Need to create an extension (FR)* *The resources at data holder level are limited, and it takes time and patience to "collect" them (NO)*
ECRIN_CRMDR	
FAIRSHARING	*There have been some test evaluations, but so far, the work has not been prioritised (NO)*
INSPIRE	
PHIRI	*There have been challenges with data harmonisation from Norway's perspective since our National Patient Registry data is so rich and detailed. By harmonising to the lowest common denominator, the Norwegian data has been vastly underutilised (NO)*

Barriers and challenges on standards enabling semantic interoperability

Standard	Comments (country)
CDISC-SDTM	
LOINC	*NPU used instead of LOINC (for the laboratory area) (DK, NO, SE, CZ)* *BfArM oversees the translation of LOINC identifiers for Germany and submits them to the Regenstrief Institute. The translation activities are supported by experts. Requests for additions or change to LOINC concepts are submitted directly to the Regenstrief Institute.The LOINC AG governed by the Advisory Board for Classifications in Health Care (KKG) advises the BMG and BfArM on questions regarding the strategic further development and application of LOINC in Germany and on the preparation of the German position for European and international committees. (DE)*

	LOINC and UCUM are made accessible through our national terminology server. LOINC is used in the domain of medical microbiology. The national 'LABCODESET' is coded using LOINC but not yet widely used.(NL) *Norway uses their own extension of the NPU database (NO)*
OMOP-CDM	
Orphanet standards	*Use of SNOMED CT instead of OrphaCodes preferred (AU)*
SNOMED CT	*Together with the NRC of Austria and Switzerland and terminological and domain experts BfArM is working on translations of SNOMED CT into German language. Change requests for extensions and modifications of SNOMED CT concepts for international and national use are evaluated by the National Release Center at BfArM. To advise on the use and further development of SNOMED CT the Federal Ministry of Health (BMG) established an Expert group under the governance of the Board of Trustees for Questions of Classification in Health Care (KKG) in 2021, in which the relevant professional groups are represented. (DE)* *Needs changes in Finnish EHR, because this is new ontology in Finland (FI)* *International Classification of Diseases for Oncology is used in pathology (HU)* *Under implementation in one of four health regions, at some areas in other regions, and in selected registries including the cancer registry (NO)* *SNOMED CT is not yet used as standard terminology in Dutch hospitals, but it is used for research projects/case studies. SNOMED CT is accessible through our national terminology server (in Dutch). The Dutch Hospital Data thesaurus is used on the national level and is mapped to SNOMED. Nictiz holds the Dutch national release centre for SNOMED. SNOMED is implemented in hospitals based on use-cases (NL)* *There is a Spanish Extension of SNOMED CT maintained by the Ministry of Health, which is the National Release Center of SNOMED CT in Spain (ES)*

Barriers and challenges on standards for interoperable communication

Standard	**Comments (country)**
DICOM	*Image exchange with IHE XDS infrastructure on national level planned (AU)*
HL7 FHIR	*Selected as prospective EHRxF standard in the Czech Republic. HL7 FHIR is envisaged in the CR for the future, however current health data repositories are typically not equipped with such interface and implementation should come gradually in upcoming years. (CZ)* *The German EHRs are defined as FHIR profiles by MIO42 and KBV. Many Domains in the German EHR are specified or will be specified as FHIR-Profiles by MIO42 and KBV in the future. (DE)* *SPMS makes available HL7-FHIR catalogues however it is not responsible for the management of the standard (PT)*
IDMP	*Z-index (G standard) contains similar data as IDMP and is aiming to reach interoperable data in agreement/alignment with IDMP (NL)*
ISO 8000-110	

8 Considerations

General considerations

- This work recognizes the evolving nature of all the standards analysed as they keep developing updates following the needs and requirements of the community of users and developers and the potential adoption and implementation of new standards such as the expected ICD-11-MMS. For the purposes of this document, the latest review on their state of the art was just after concluding the interviews and having received feedback from the promoters of the standards in the synthesis of the outputs (October 2022). It is also worth highlighting that the inclusion of new standards was closed in May 2022. Along with the analysis of the country surveys we have come up with some standards that have not been considered in this document, but may well deserve further attention in a new version (e.g. openEHR in the case of Spain, or NPU in the case of Norway, Sweden, Denmark, and the Czech Republic).

- Controlled vocabularies widely used for recording clinical conditions and procedures in the process of care, specifically ICD or ICPC families, have been intendedly omitted from this analysis as the aim of this report with regard to this issue was providing a general solution for the semantic interoperability of clinical and procedural concepts. By adopting this approach, we are not suggesting by any means that the way clinical conditions and care procedures are usually recorded at origin (i.e. in the place of treatment) should be changed for these other standards. Rather, our vision is that, irrespective of the controlled vocabulary used, once data is made available for secondary purposes, data processors, instead of providing multiple mapping (e.g. ICD-X to ICD-Y editions to ICPC-Z) could benefit from the use of comprehensive semantic standards mapped out to many controlled vocabularies.

- In addition, we have not considered formatting standards intended to map relationships among data in the form of triplets or *subject-predicate-object* structures such as RDF, OWL, SKOS, etc. These W3C standards may be required to support applications querying data at a later step, for example, while configuring terminology servers mapping across standards or controlled vocabularies, or mapping between similar variables and values across different data sources. Their use and implementation have been deemed out of the scope of this report.

- In a similar way, we have not considered other metadata standards, such as DDI-RDF Discovery Vocabulary (Disco) aimed at publishing metadata on research and survey datasets, more specific to the research domains. However, we consider that their use may be integrated with DCAT in providing a standard scheme for information about the logical structure (variables, concepts, etc.) of tabular datasets.

- We have to acknowledge that some care domains have not been covered in this report; in particular, functional status and disabilities, and nursing care and other social aspects of care. For the former, it may be worth deepening the potential of the International Classification of Functioning, Disability and Health (WHO-ICF). In the latter case, there are two ontologies that may result of interest: the ontology NANDA-NIC-NOC or the one in the project NursingOntos, developed in the University of Porto in Portugal.

- Likewise, we have also to acknowledge that there are some systems of concepts supporting the standardised description of continuity of care within a health system across healthcares services and providers (e.g., ContSys - EN ISO 13940:2015), which imply a harmonised description of healthcare clinical or administrative information systems. Although this type of systems of concepts are out of the scope of this report, they could be of interest when describing the provenance of data.

- In a similar fashion, this report does not include the assessment of standards aiming the harmonisation of particular features of health data sources such as data quality (e.g. W3C-Data Quality Vocabulary (DQV) or the data usage schemas (i.e. Data Usage Ontology (DUO), MPEG-21 REL(ISO/IEC 21000-5), Open Digital Rights Language (ODRL), or XACML policy (Xtensible Access Control Markup Language)). These standards are not included given that they are not being widely implemented in the domain of health.

- Other existing standards such as the Generic Statistical Information Model (GSIM), referencing types of information used in the production of official statistics, or the ISO/IEC 11179 Metadata Registry (MDR), representing the organisation of metadata registries, were not considered either, as are they are out of the scope of this report.

- Finally, out of the analyses of the results some criteria have been deemed as more relevant than others in the elaboration of the conclusions and recommendations; thus:

    - How well the standard fits the core principles of CAMSS, in particular openness
    - How well the standard fits Interoperability Layers as in EIF
    - Whether the standard is able to map existing controlled vocabularies
    - Whether the standard has looked for joint development with others
    - How wide the experience on the standard has been
    - Whether the standard is widely implemented/is reported to be widely implemented

    -

Considerations on the technical support of these standards

As aforementioned, the technological interoperability of the analysed standards is assumed based on the availability of technical documentation for their implementation and their wide

adoption within their specific domains. Still, there is some work to be done with WP7 on the interface between the semantic and syntactic communication standards in this report and the work they are doing with regard to the communication between HealthData@EU nodes. In particular, WP7 is yet to decide whether the content of the data access application or data request will include low-level (i.e. precise information) thus requiring a standard to catalogue the data source at meta-data or data dictionary level. Once the decision has been made, the conclusions on communication standards in this report may require some adaptation. It is also worth commenting that the architecture of the HealthData@EU is still under discussion. Whether the final architecture may influence any of the conclusions and recommendations of this report is unknown.

Considerations on the governance and sustainability

The technical analysis developed on the standards is insufficient when it has to translate into a recommendation on its use, particularly if this recommendation has an effect at European level. For this reason, the conclusions of this document necessarily have to be complemented with the insight gathered in other TEHDAS work packages, in particular, WP4, WP5 and the upcoming work in WP6.

In the case of Governance, further developments are needed, partly provided by the latest WP5 deliverable, to determine who (Institutions in the HealthData@EU) should be responsible for the implementation of the standards and its maintenance. On a different line, the upcoming WP6 thread on the quality and utility labelling and an associated maturity model raises governance elements on, for example, who will be implementing the model, who will be responsible for the assessment, and who will be in charge of the supervision and promotion of the datasets to a higher level of utility. Linked to this, whether a model of incentives for improvement should be implemented and supervised.

In addition, the implementation of those recommended standards would require an analysis of the costs of implementation and maintenance, and how these costs should be financed. From WP4 some insight is expected in this respect.

9 Highlights and recommendations

The objective of this scoping review was the identification and assessment of standards of interoperability and the provision of guidance on their potential use in the context of HealthData@EU data life cycle. As a result, 9 standards for data discoverability, 5 standards enabling semantic interoperability, and 4 standards for interoperable communication across data institutions were analysed.

With some caveats, the methodology of assessment (CAMMS) has been proven suitable for a formal assessment of those standards in the context of the development of the

HealthData@EU. The assessment has yielded a notion of the pros and cons of an eventual use of the standards at different stages of the data life cycle for secondary use (i.e., data discovery, semantic interoperability and interoperable communications).

In the following paragraphs, the main findings and conclusions will be highlighted, paving the way for a number of recommendations.

Standards for data discoverability

Overall conclusions

Overall, among the studied standards, DCAT-AP and INSPIRE were considered best equipped for data discoverability, as per CAMMS evaluation.

These standards, developed in the context of OpenData and Public Administration, are however not so well equipped to respond to the needs of the research communities. On the contrary, those standards developed in the context of research communities, that ranked lower in reusability or in the assessment of effectiveness and efficiency, ranked top in the user-centric approach which highlights the actual relevance of these standards in data discoverability for specific research communities - i.e. data types as biosamples, gene sequences, clinical research data, medical records etc.

None of the above standards are part of a national framework or national strategy in the context of health data or have been widely adopted as a standard for discoverability.

Importantly, DCAT-AP, INSPIRE and CESSDA have developed some mutual collaboration.

On the core principles

- Those standards ranking the highest were CESSDA-CMM (94%), DCAT-AP2 (100%) and INSPIRE (100%). The rest ranked over 70%, openness and technological neutrality and data portability being their main features.

  Specifically, those standards ranking the highest in Openness were BBMRI-MIABIS, CESDDA-CMM, DCAT-AP2 and INSPIRE; those standards ranking the highest in Transparency were BEACON, BBMRI-MIABIS, DCAT-AP2, FAIRSHARING and INSPIRE; those standards ranking the highest in Reusability were CESSDA-CMM, DCAT-AP2, FAIRSHARING, and INSPIRE; and, those standards ranking the highest in Technological neutrality and data portability were CESSDA-CMM, ECRIN-CRMDR, INSPIRE and PHIRI.

On the principles of interoperability

- Those standards ranking the highest in the layer of interoperability were BBMRI-MIABIS and CESSDA-CMM (96%) and DCAT-AP2 and INSPIRE (100%).

  Specifically, those standards ranking the highest in Governance were DCAT-AP2 and INSPIRE (100%); all the standards, except Beacon and PHIRI-HIP ranked the highest in Organisational interoperability; all the standards, except PHIRI-HIP, ranked the highest in Semantic interoperability; and all the standards ranked the highest in Technological interoperability.

RECOMMENDATION 1: In HealthData@EU, data discoverability may benefit from the combined use of generic standards and domain-specific standards.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 11 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 2: This combined use may on the side of data preparatory bodies require the implementation of a two-step process supporting the phase of data discovery; a) a first step focusing on gathering high-level knowledge on the data sets available that is agnostic to the domain or the type of data; and, b) a second step where the focus is the actual content of the data source, that would be domain- data type-specific.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 8 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 3: Before an eventual adoption of a generalistic standard, there may be a need for a discussion on whether more detailed information is required when describing health data sources (eg, coverage, provenance), information that is agnostic to the domain or data types.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 11 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 4: Before an eventual adoption, there may be a need for domain-specific standards to improve the governance of interoperability, preparing the standard to be mapped to EIRA (European Interoperability Reference Architecture) and implementing a mechanism to assess conformity in the implementation.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 10 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

Standards enabling semantic interoperability

Overall conclusions

All the standards analysed got a similar interoperability overall score (around 80%), ranging between 74% in Orphanet standards and 82% in OMOP-CDM.

The adoption of these standards is quite uneven. While in the case of SNOMED CT and LOINC there is a wide experience across Europe, the adoption of Orphanet standards is more limited, and the adoption of OMOP-CDM is essentially linked to research projects on specific domains, although ranking the highest in the principles for cooperation among public institutions. CDISC-STDM is not used in any of the countries that provided insight.

Importantly, there are mappings between SNOMED CT to Orphanet standards (ORPHAcodes), OMOP-CDM, CDISC-SDTM, and LOINC has reached a collaboration agreement with SNOMED CT. In addition, recent editions of ICD (except from ICD-11-MS) and ICD-O are mapped to SNOMED CT, as well as, are being used in the Human Phenotype Ontology and in the GLobal Alliance for Genomics and Health. Finally, SNOMED CT is ISO-IDMP compliant allowing the extension of EMA case safety reports.

On the core principles

- All the standards in this group scored around 70% in the core principles - ranging from 66% in Orphanet standards and 75% in LOINC. These mild figures are a reflection of the low scores the five standards got in reusability.

  Indeed, as a semantic layer they are very well equipped to address the specificities of their respective ontologies (SNOMED CT, LOINC and Orphanet standards) or common data models (OMOP-CDM and CDISC-SDTM) being difficult for them to expand out to other semantic domains. As a consequence, none of the standards are individually able to fully cover the need of semantic standardisation across all potential data collections.

On the principles of interoperability

- The five standards showed mild scores according to the layers of the European interoperability framework (overall score between 74% and 82%). SNOMED CT, LOINC and Orphanet standardsgot the highest scores.

Importantly, this score reflects the low score obtained in the item “governance of interoperability” as none could be mapped to European Interoperability Reference Architecture (EIRA) and just those aforementioned have already been recommended and catalogued as an EU standard or being supported by an EU member as a standard of interest.

Except SNOMED CT and Orphanet standards, despite the intrinsic semantic value of the rest of the standards, they showed intrinsic difficulties in the cross-country reuse of the derived outputs.

RECOMMENDATION 5: Although none of the above standards can cover all the data types of interest in the HealthData@EU, data preparatory institutions (i.e. those with the role of data holders) could safely use them as a semantic layer when standardising their data. As per CAMMS assessment SNOMED CT has been shown to be the best equipped ontology to cover semantic interoperability across controlled vocabularies and taxonomies referred to medical concepts.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 11 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 6: As the current mapping of medical concepts from taxonomies and controlled vocabularies to SNOMED CT is not fully completed, we recommend the European Commission fostering this effort and the Member States to progressively deploy SNOMED CT as an ontology of reference for medical concepts.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 12 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 7: There will still be a need for development and sharing of semantic maps to other than medical concepts as concepts from other determinants of health (i.e., social, cultural and economic determinants, environmental determinants, genetic determinants). As these concepts are often instrumental to specific uses or research projects, we would recommend data holders to enrich their data collections with these maps and systematically share with other data holders within HealthData@EU.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 11 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

RECOMMENDATION 8: The preparation of data for secondary use should not be limited to the mapping of concepts, but to the development of data models, considering as entities in

the model: individuals, place of residence, place of treatment, contacts with the system, treatments, and time. As for this purpose, we do recommend the European Commission and Member States to design and implement a specific development in this respect, taking as an inspiration how the initiative fostering OMOP-CDM has addressed openness, transparency, technological neutrality and data portability, and cooperation among public institutions.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 11 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

Standards meant interoperable communication

Overall conclusions

- Overall, except for IDMP, the standards of communication scored 90% and over in accordance with the CAMSS assessment showing suitable for semantic and syntactic communication within the HealthData@EU.

  DICOM and HL7-FIHR both scored 99% while ISO-8000-110 reached 96% reflecting a lower score in principles for cooperation across institutions. IDMP low figures reflect issues on transparency, reusability, security and privacy and a lack of assessment of effectiveness and efficiency.

  As per the adoption, DICOM has been widely adopted in all the countries surveyed in this study, and HL7-FHIR to a considerable extent. This is not the case for ISO 8000-110 just adopted in one institution of the surveyed countries. IDMP is used in the communication between manufacturers and EMA but it is not mentioned at country level.

  Importantly both DICOM and HL7-FHIR have set up a mutual collaboration and with the abovementioned semantic standards. IDMP has focused on SNOMED CT and HL7-FHIR. The level of cooperation is uneven. Finally, DICOM has established cooperation with the Cancer Genome Atlas program in the USA.

  We do have to recognize that there may be domain-specific communication standards that have not been analysed in this report and may well serve the needs of specific communities.

On the core principles

- DICOM and HL7-FIHR scored 100% in the principles of openness, transparency, reusability and technological neutrality and data portability. The slightly lower score in ISO 8000-110 was in openness as it is not possible for everyone to study standard and specifications with no restrictions.

On the principles of interoperability

- All standards suffer from some issue in the item “governance of interoperability”. None of them can be mapped to the European Interoperability Reference Architecture. In addition, IDMP suffers from the lack of mechanisms to assess conformity of the implementation.

RECOMMENDATION 9: According to the CAMMS assessment DICOM and HL7-FHIR have been proven the best equipped standards of communication. We do recommend the European Commission and the Member States to assess the implementation efforts and technical difficulties for the above to be the communication standards between the health data access bodies, for example, in the communication between secure process environments.

WP6 vote on the recommendations: out of the 13 respondents from WP6, 12 scored 7 or higher. In annex IV the responses gathered from the different countries participating in the consultation may be found.

10 Annexes

I. CAMSS outputs

II. Standards fiches

III. Country surveys outputs

IV. Recommendations to enhance interoperability at HealthData@EU - Consultation to WP6 partners

Annex I - CAMSS outputs

CAMSS assessment summary outputs can be accessed and commented on this link.

Annex II - Standards fiches

Standards' fiches can be accessed and commented on in this link.

Annex III - Country survey outputs

Experience within the countries

Experience on standards for data discoverability

Standard	Beacon	BBMRI-MIABIS	Bio-image archive	CESSDA CMM	DCAT-AP	ECRIN CRMDR	FAIRSHARING	INSPIRE	PHIRI
Austria	No	Yes	No	No	No	No	No	No	No
Czech_Republic	No	Yes	Yes	Yes	Marginally	Yes	Yes	Yes	No
Denmark	Yes	No	No	No	No	No	No	No	No
France_AMU	No	No	No	No	No	No	No	No	No
France_ANS	-	-	-	-	-	-	-	-	-
France_HCL	No	No	No	No	No	Yes	No	No	No
France_HDH	-	-	-	-	Yes	-	-	-	-
France_Orphanet_Inserm	Yes	Yes	-	-	Yes	Yes	Yes	-	No
Finland	Yes	Yes	No	Yes	Yes	No		Yes	Yes
Germany	No	No	No	No	No	No	No	No	No
Hungary	No	No	No	No	No	No	No	No	No
Hungary_Semmeelweis	Yes	Yes	No	Yes	No	Yes	No	No	Yes
Ireland	No	No	No	No	No	No	No	No	No
Netherlands	Yes (It is believed)	Yes	No	No	No	No	Average	No	No
Norway_1	No+	Yes	Yes	Yes	Yes	Yes	No	Yes-	Yes

Norway_2	No+	Yes	No	No	Yes	No	No	Yes-	No
Portugal	No	No	No	No	No	No	No	No	No
Spain	No	No	No	No	No	No	No	No	No
Sweden	No	No	No	No	Yes	No	No	No	No

Experience on standards meant to build data models

Standard	CDISC SDTM	LOINC	OMOP CDM	Orphanet standards	SNOMED CT
Austria	Yes	Yes	Yes	Yes	Yes
Czech_Republic	No	Yes	No	Yes	Yes
Denmark	No	No	No	No	Yes
France_AMU	Know the main principles of standard	Know the main principles of standard	Know the main principles of standard	No	Yes (education purpose)
France_ANS	-	-	Yes	-	-
France_HCL	No	Yes	Yes	No	Yes
France_HDH	-	Yes	Yes	-	Yes
France_Orphanet_Inserm	Yes	Yes	Yes	Yes	Yes
Finland	No	Yes	Yes	Yes	Yes
Germany	-	Yes	Yes	Yes	Yes
Hungary	No	Yes	No	Yes	Yes
Hungary_Semmeelweis	No	Yes	No	Yes	Yes
Ireland	No	Yes	No	No	Yes
Netherlands	Only superficially	Yes	Only superficially	Yes	Yes
Norway_1	No	No+	Yes	Yes	Yes
Norway_2	No	No+	Yes	Yes-	Yes
Portugal	No	Yes	No	Yes	Yes
Spain	No	Yes	Yes	Yes	Yes
Sweden	Yes	Yes	Yes	Yes	Yes

Experience on standards enabling interoperable communication

Standard	DICOM	HL7 FHIR	IDMP (SPOR)	ISO 8000-110
Austria	Yes	Yes	Yes	No
Czech_Republic	Yes	Yes	Yes	No
Denmark	Yes	Yes	Yes	No (not offically listed for usage)
France_AMU	Know the main principles of standard	No	No	No
France_ANS	Yes	Yes	-	-
France_HCL	Yes	Yes	No	No
France_HDH	Yes	-	-	-
France_Orphanet_Inserm	No	Yes	No	Yes
Finland	Yes	Yes	?	No
Germany	Yes	Yes	Yes	No
Hungary	Yes	Yes	No	No
Hungary_Semmeelweis	Yes	Yes	No	No
Ireland	Yes	Yes	Yes	No
Netherlands	Yes	Yes	Yes	-
Norway_1	Yes	Yes	Yes	No
Norway_2	Yes	Yes	Yes-	No
Portugal	Yes	Yes	Yes	No
Spain	Yes	Yes	Yes	No
Sweden	Yes	Yes	No	No

Adoption and inclusion in the National Framework

Experience on standards for data discoverability

Standard	Beacon	BBMRI-MIABIS	Bio-image archive	CESSDA CMM	DCAT-AP	ECRIN-CRMDR	FAIRSHARING	INSPIRE	PHIRI
Austria	No	Only in Biobanks	No	No	No	No	No	No	No
Czech_Republic	No	Yes	No	No	No	Partially	No	Partially	No
Denmark	No	No	No	No	No	No	No	No	No
France_AMU	-	-	-	-	-	-	-	-	-
France_ANS	-	-	-	-	-	-	-	-	-
France_HCL	No	No	No	No	No	Yes	No	No	No
France_HDH	-	-	-	-	No	-	-	-	-
France_Orphanet_Inserm	Yes	Yes	-	-	Yes	Yes	Yes	-	-
Finland	In use in THL biobank	In use in THL Biobank for cohort descriptions that are available for both national and international	No	Data archived in Finnish Social Science Data Archive is described in CESSDA Data Catalogue using CMM.	Doesn't seem that it's yet really widely used. The Opendata.fi portal has its own DCAT-AP extension.	-	-	In use in Finland but the precise situation is unclear.	In use but not widely. The most important Covid-19 data from THL and Finland

		users. FinBB's Ingenious Cohort service descriptions are also based on MIABIS. Terminology is adopted in biobanking at European level.							has been described using PHIRI metadata specification in the European Health Information Portal.
Germany	-	-	-	-	-	-	-	-	-
Hungary	No	No	No	No	No	No	No	No	No
Hungary_Semmeelweis	No	No	No	No	No	No	No	No	No
Ireland	No	No	No	No	No	No	No	No	No
Netherlands	-	-	-	-	-	-	?	-	-
Norway_1	No, but someone in Oslo is actively promoting Beacon.	-	The Norwegian University of Life Sciences uses this BioImage meta-data standard for	Norwegian Agency for Shared Services in Education and Research (Sikt) is a CESSDA ERIC consortium member. Within CESSDA ERIC, Sikt provides	The Norwegian Directorate of eHealth has developed metadata specification more or less based on DCAT	ECRIN currently has seven Member Countries (France, Germany, Hungary, Italy, Norway, ...)	The Norwegian Directorate of eHealth has developed a FAIR guideline and evaluation tool for	Norway has been a part of the Inspire "community" throuh the Norwegian Digitalization Agency and the	The Norwegian Directorate of Health and the Norwegian Institute for Public

			academic and research purposes.	ongoing training, webinars,workshops on CESSDA strategy, expertise, tools, development and impact.	properties. All health data sources in Norway has to share their metadata according to this specification		health data sources (registries)	Norwegian Mapping Authority. All sector are sharing metadata with the National Common Data Catalogue. INSPIRE is not well known in the Healthcare sector	Health are partners in the PHIRI project that is laying the foundation to build a European Research Infrastructure on Population Health to be used to overcome future crises (e.g. COVID-19)
Norway_2	-	-	-	-	The Norwegian Directorate of	Portugal and Spain) and two Observer	The Norwegian Directorate	Norway has been a part of the	-

					eHealth has developed metadata specification more or less based on DCAT properties. All helath data sources i Norway has to share their metadata according to this spescification	Countries (Czech Republic and Switzerland).	of eHealth has developed a FAIR guildeline and evaluation tool for healtdata sources (registries)	Inspire "community" throuh the Norwegian Digitalizatio n Agency and the Norwegian Mapping Authority. All sector are sharing metadata with the National Common Data Catalogue. INSPIRE is not well known i the Helathcare sector	
Portugal	No	No	No	No	No	No	No	No	No
Spain	No	No	No	No	No	No	No	-	-
Sweden	No	No	No	No	Yes	No	No	No	No

Experience on standards enabling semantic interoperability

Standard	CDISC SDTM	LOINC	OMOP CDM	Orphanet standards	SNOMED CT
Austria	No	Yes	No	No	Partly, several value sets for some applications
Czech_Republic	No	No	No	No	No
Denmark	No	No	No	No	Yes
France_AMU	-	-	-	-	In process of becoming a member country of IHTSDO
France_ANS	-	Yes	Yes	-	-
France_HCL	No	Yes	Yes	No	Yes
France_HDH	-	Yes	Yes	-	Yes
France_Orphanet_Inserm	-	Yes	Yes	Yes	Yes
Finland	Yes	No	Yes	Partly implemented or plan to implement in EHR and HILMO registers.	Partly, especially in pathology medicine
Germany	-	Yes	-	Yes	Yes (basic terminology for patient record and Medicinal Information Objects)
Hungary	No	Yes	No	No	No
Hungary_Semmelweis	No	No	No	-	No
Ireland	No	Partly	No	No	Partly - All national maternity hospitals. Implemented throughout St. James

					hospital - large national teaching hospital is an exemplar for SCT.
Netherlands	Low	Yes	Low	??	Average
Norway_1	-	No+	The Norwegian Cancer Registry and the University of Oslo are datapartners in the EDEN project. The Norwegian Directorate of Health aim to participate. Some of the IT tools are downloaded. A subtask in the Nordic Commons?	No+	No+/Yes-
Norway_2	-	-	The Norwegian Cancer Registry and the University of Oslo are datapartners in the EDEN project. The Norwegian Directorate of eHealt aim to participate. Some of the IT tools are downloaded. A subtask in the Nordic Commons?	-	No+
Portugal	No	Yes	No	Yes	Yes
Spain	-	It is used in most Autonomous Communities from Spain	No	The Ministry of Health provides a mapping from SNOMED CT a ORPHA. We do not have usage statistics.	Yes, SNOMED CT is used in almost every Autonomous Communities
Sweden	No	No	No	No	Yes

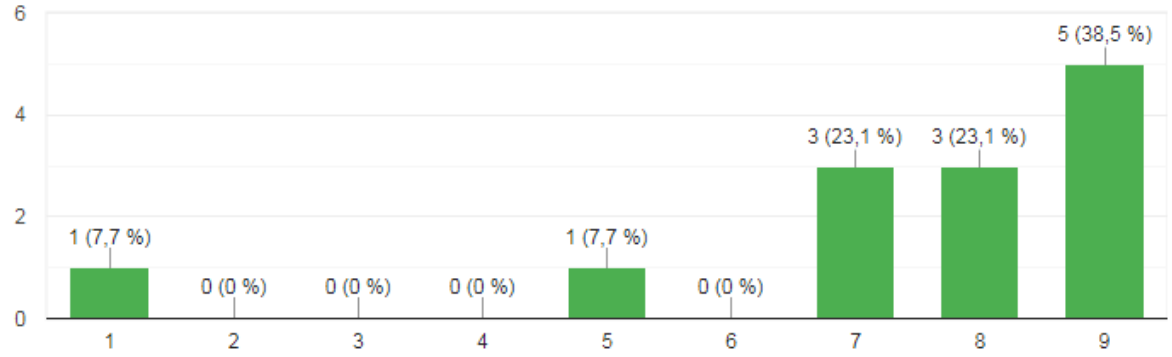
Experience on standards enabling interoperable communication

Standard	DICOM	HL7 FHIR	IDMP (SPOR)	ISO 8000-110
Austria	Yes	No yet	No yet	-
Czech_Republic	Yes	No	No	No
Denmark	Yes	Yes	Yes	No (not officially listed for usage)
France_AMU	-	-	-	-
France_ANS	Yes	Yes	-	-
France_HCL	Yes	Yes	-	No
France_HDH	Yes	-	-	-
France_Orphanet_Inserm	Yes	-	-	Yes
Finland	Yes	Yes	No yet	No
Germany	Yes	Yes	Yes	-
Hungary	Yes	No	No	No
Hungary_Semmeelweis	Yes	Yes	No	No
Ireland	Yes	No	Planned	No
Netherlands	Yes	High	No	-
Norway_1	This is the major standard for digital medical applications handling medical imaging and was established in 1992 . All hospitals in Norway use DICOM for medical image	No, but increasing. The Norwegian Directorate of e-health has issued a high-level recommendation to use HL7 FHIR for integrations based on data sharing in the healthcare	Not yet, but will be.* The Norwegian Directorate of e-health recommends the use of IDMP for describing product-specific information. The use of Medicinal Product Identifier	No

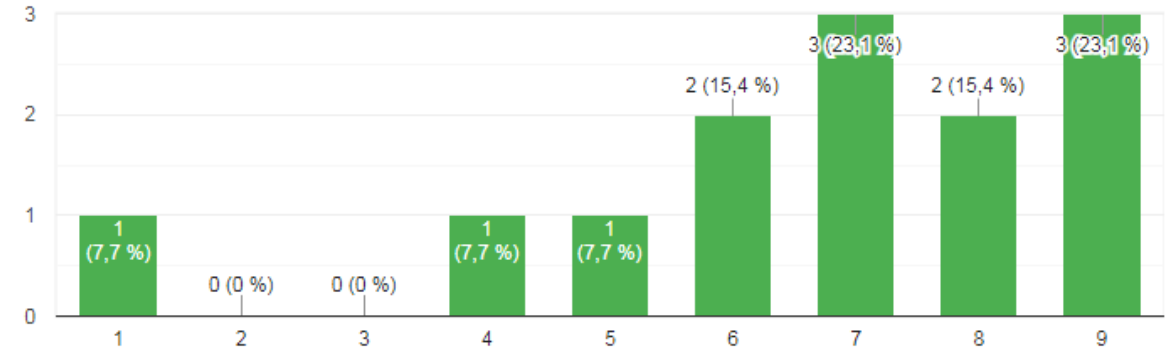
	communication.	sector in 2019. The Directorate also recommends using SMART on FHIR for integration of applications to EHRs.	is recommended throughout the value chain for medicinal products.	
Norway_2	-	No, but increasing. The Norwegian Directorate of e-health has issued a high-level recommendation to use HL7 FHIR for integrations based on data sharing in the healthcare sector in 2019. The Directorate also recommends using SMART on FHIR for integration of applications to EHRs.	Not yet, but will be.* The Norwegian Directorate of e-health recommends the use of IDMP for describing product-specific information. The use of Medicinal Product Identifier is recommended throughout the value chain for medicinal products.	No
Portugal	Yes	Yes	No	No
Spain	DICOM is widely used.	No, HL7 CDA is used for Patient Summary, but not in national or local systems. HL7 FHIR is not widely adopted. OpenEHR is being considered as a complement	No, National codes are used. For cross-border use, these codes are translated into English.	No
Sweden	Yes	No	No	No

Annex IV - Recommendations to enhance interoperability at HealthData@EU - Consultation to WP6 partners

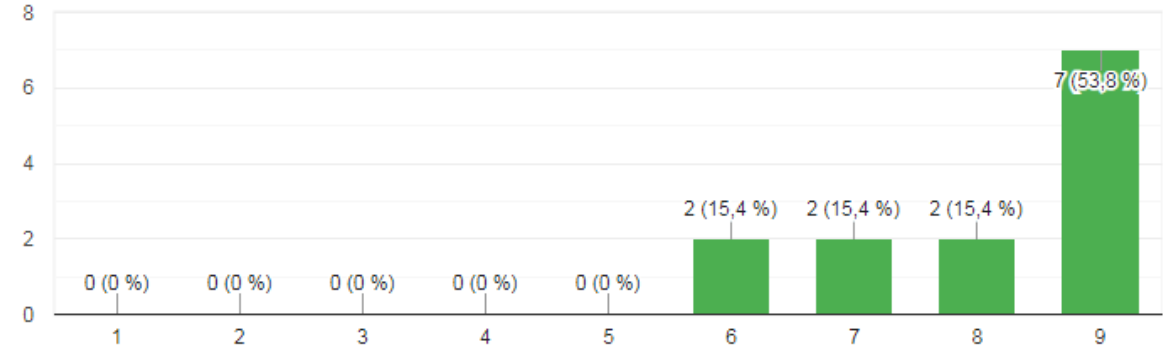
RECOMMENDATION 1



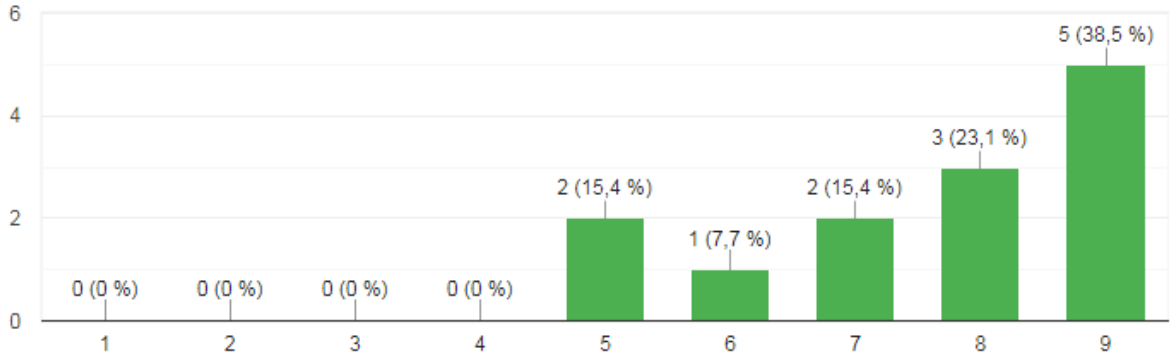
RECOMMENDATION 2



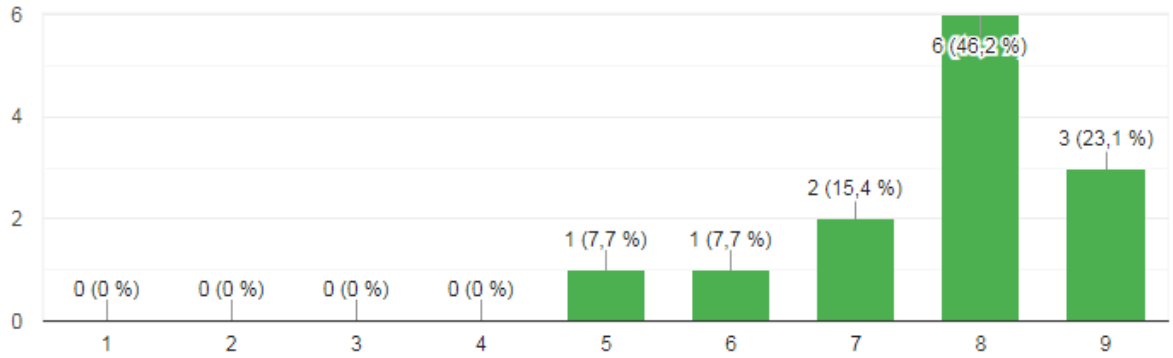
RECOMMENDATION 3



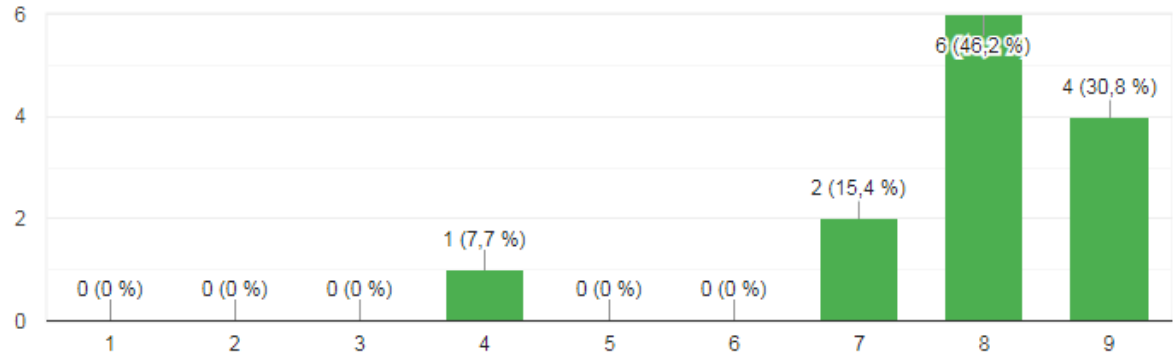
RECOMMENDATION 4



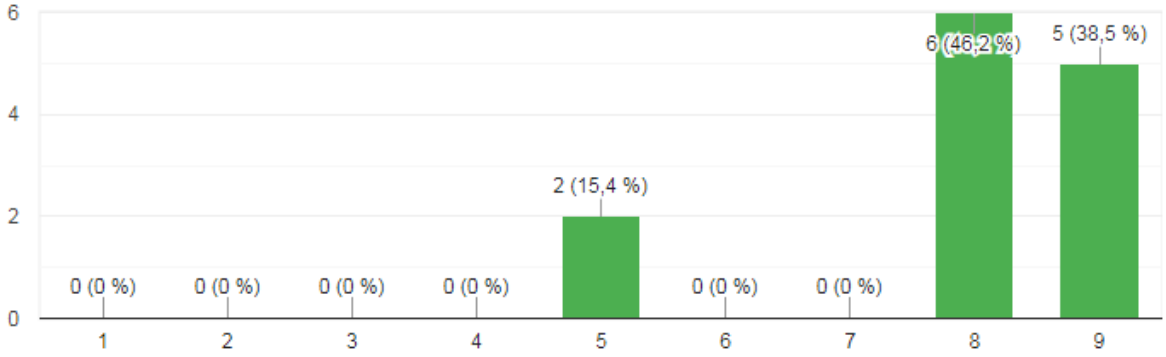
RECOMMENDATION 5



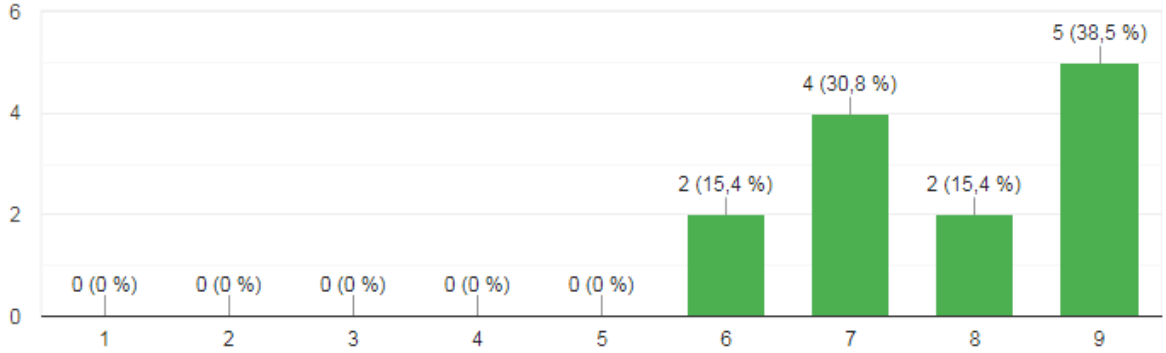
RECOMMENDATION 6



RECOMMENDATION 7



RECOMMENDATION 8



RECOMMENDATION 9