Improving cancer diagnosis and prediction with AI and big data

A Multimodal AI-based Toolbox and an Interoperable Health Imaging Repository for the Empowerment of Imaging Analysis related to the Diagnosis, Prediction and Follow-up of Cancer

Deliverable 3.4

INCISIVE Standardization Suggestions

(Final Version)

WP 3 – Standardization Actions

28-07-2023

Revision 1.0

Status: Final Version

Grant Agreement n 952179

DOCUMENT CONTROL	
Project reference	Grant Agreement number: 952179
Document name	INCISIVE Standardization Suggestions (Final Version)
Work Package	WP 3
Work Package Title	Standardization Actions
Dissemination level	PU
Revision	1.0
Status	Final Version
Reviewers	AUTH (Ioanna Chouvarda) and CeRICT (Salvatore D'Antonio, Giovanni Mazzeo)
Beneficiary(ies)	

Dissemination level:
PU = Public, for wide dissemination (public deliverables shall be of a professional standard in a form suitable for print or electronic publication) or CO = Confidential, limited to project participants and European Commission.

AUTHORS		
	Name	**Organisation**
Document leader	Sara Martínez Alabart	FTSS
	Susanna Aussó Trias	FTSS
Participants	Sara Martínez Alabart	FTSS
	Shulei Huang	FTSS
	Zisis Sakellariou	CERTH
	Tasos Bachtses	CERTH
	Alexandra Kosvyra	AUTH
	Caroline Barelle	ED
	Alberto Gutierrez-Torre	BSC
	Vladimir Petrovic	VIS
	Magdalena Kogut-Czarkowska	TLX
	Salvatore D'Antonio	CeRICT
	Pablo Mezzon	ED

REVISION HISTORY				
Revision	**Date**	**Author**	**Organisation**	**Description**
0.0	08.06.2023	Sara Martínez Alabart	FTSS	Initial version of the document
0.0	26.06.2023	Sara Martínez Alabart	FTSS	New index of contents
0.0	27.06.2023	Sara Martínez Alabart	FTSS	Integration of content regarding sections 1, 2, 3, 4, 5, 6 and 7
0.0	03.07.2023	Alexandra Kosvyra	AUTH	Integration of content regarding sections 4.2 and 4.7.5.4
0.0	07.07.2023	Magdalena Kogut-Czarkowska	TLX	Integration of content regarding section 3.2.7.1, 4.4 and 4.7.1
0.0	13.07.2023	Tasos Bachtses	CERTH	Integration of content regarding sections 4.7.4.3
0.0	14.07.2023	Zisis Sakellariou	CERTH	Integration of content regarding sections 4.7.4.1 and 4.7.5.4
0.0	14.07.2023	Alberto Gutierrez-Torre	BSC	Integration of content regarding sections 4.7.4.3
0.0	14.07.2023	Caroline Barelle	ED	Integration of content regarding sections 4.7.5.1, 4.7.5.3 and 4.7.5.4
0.1	14.07.2023	Sara Martínez Alabart	FTSS	First Draft, Integration of content regarding sections 1, 2, 3, 4, 5, 6 and 7
0.2	17.07.2023	Vladimir Petrovic	VIS	Integration of content regarding sections 4.7.1.1
0.2	18.07.2023	Pablo Mezzon	ED	Integration of content regarding sections 4.7.5 and 4.7.5.4
0.2	19.07.2023	Salvatore D'Antonio	CeRICT	Integration of content regarding sections 3.2.7.1
0.2	19.07.2023	Sara Martínez Alabart	FTSS	Document ready for internal review Final draft
0.3	19.07.2023 - 27.07.2023	Sara Martínez Alabart, Shulei Huang	FTSS	Internal review and addressing of comments completed
1.0	28.07.2023	Sara Martínez Alabart Shulei Huang	FTSS	Document ready for submission

Disclaimer and statement of originality

The content of this deliverable represents the views of the authors only and is their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of Contents

Table of Figures

Table of Tables

Abbreviations and Dictionary

Abbreviation	Description
AI	Artificial Intelligence
API	Application Programming Interface
ATC	Anatomical Therapeutic Chemical
ATNA IHE	The Audit Trail and Node Authentication (ATNA) Integration Profile
CDM	Common Data Model
CDR	Clinical Data Repository
CDS	Clinical Decision Support
CEF	Connecting Europe Facility
CN	Central Node
CPT	Current Procedural Terminology
CRT	Chemoradiotherapy
CT	Computerized axial Tomography
CVX	Vaccine administered code set
DBB	Databases
DGA	Data Governance Act
DICOM	Digital Imaging and Communication In Medicine
DIMSE	DICOM Service Elements
DP	Data provider
EHDS	European Health Data Space Regulation
eHealth EIF	European eHealth Interoperability Framework
EHR	Electronic Health Record
EMDN	European Medical Device Nomenclature
EMR	Electronic Medical Record
ETL	Extract-Transform-Load
EU	European Union
EUDAMED	European Database on Medical Devices
FDA	Food and Drug Administration
FN	Federated Node
FS	Federated Space

GDPR	General Data Protection Regulation
HIS	Hospital Information System
HL7 CDA	Clinical Document Architecture: Standard of Health Level Seven Organization
HL7 FHIR	Fast Healthcare Interoperability Resources: Standard of Health Level Seven Organization
HL7 v2 or v3	Health Level Seven Standard Version 2 or Version 3
HL7	Health Level Seven: International not-for-profit organization for standards developing dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing, and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services.
HPO	The Human Phenotype Ontology
HTTP	Hypertext Transfer Protocol
IBM	International Business Machines Corporation
ICD	International Statistical Classification of Diseases and Related Health Problems
ICT	Information and Communications Technology
IEEE Devices	IEEE Devices Numbers
IHE	Integrating the Healthcare Enterprise
IOD	Information Object Definition
ISO	International Organization for Standardization
IT Infrastructure	Information Technology Infrastructure
IVDR	In Vitro Diagnostic medical devices Regulation
JSON	JavaScript Object Notation
LIS	Laboratory Information System
LOINC	Logical Observation Identifiers Names and Codes
MaaS	Model-as-a-Service
MDC	Medical Device Certification
MDR	Medical Devices Regulation
MEDCIN	Medicom Systems Terminology
MED-RT	Medication Reference Terminology
MRI	Magnetic Resonance Imaging
NANDA	North American Nursing Diagnosis Association
NDC	FDA's National Drug Code
NDDR	Nearly raw raster data
NDF-RT	National Drug File Reference Terminology

NHS	National Health Service
NIFTI	Neuroimaging Informatics Technology Initiative
OASIS	Organization for the Advancement of Structured Information Standards
OAuth	Open Authorization protocol
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OpenEHR	Open Electronic Health Record
PACS	Picture Archive and Communication System
PDF	Portable Document Format
RAD	Radiology
ReEIF	Refined eHealth European Interoperability Framework
REST	Representational state transfer
RIS	Radiology Information System
SNOMED CT	Systematized Nomenclature of Medicine-Clinical Terms
SOA	Service-Oriented Architecture
SOP	DICOM Service-Object Pair
SR	DICOM Structured Report
UCUM	Unified Code for Units of Measure
UDI-DI	Unique Device Identifier – Device Identifier
UI	User Interface
UMLS	Unified Medical Language System
UNII	Unique Ingredient Identifier
WHO	World Health Organisation
WP	Work Package
XDS	Cross Enterprise Document Sharing
XML	eXtensible Markup Language

1. Presentation

1.1. Document Purpose

This deliverable includes recommendations for any worldwide project or initiative that wants to operate with global standards.

It presents the INCISIVE interoperability framework to show the methodology from the beginning, the use of global standards, and the specification and definition used to be an example for other initiative or research as INCISIVE platform regarding standard exchange of clinical data and medical images.

It also explains the issues, recommendations, possible future steps, and the possibility of reutilization of the implementation guides defined, and the advantages of using global standards.

1.2. Document Scope

The scope of this deliverable covers all aspects covered within INCISIVE project regarding interoperability. Suggestions for standardization in the health sector are also introduced.

It also describes how was defined the interoperability framework, what was the methodology that INCISIVE use to analyse the identification process and use cases, explains how was done the specification of the different interoperability scenarios, indicates the importance of explaining the potential risks and maintenance plans and how this was done by INCISIVE, provides a comprehensive overview about interoperability frameworks and recommendations about standards suggestions for the future projects, and concludes with the importance of describing the release process and the reuse of the interoperability frameworks.

1.3. Related Documents

INCISIVE_Interoperability_Framework document.

2. Introduction

This section aims to make an introduction to health interoperability, its layers, and the health standards of each layer. It also gives a specific approach for European or international projects while making recommendations on how to define an interoperability framework.

2.1. Interoperability and Standards in Healthcare

To understand the interoperability, it is necessary to know the difference between information and data. Data is a collection of facts, while information puts those facts into context. While data is raw and unorganized, information is organized. Data points are individual and sometimes unrelated. Information maps out that data to provide a big-picture view of how it all fits together.

Interoperability refers to the capacity of different systems, devices, or applications to exchange information without altering its meaning. The implementation of interoperability is crucial to resolve data integration issues between systems, mitigate data inconsistencies, and enable effective management of potential changes. It is designed to convert information from one system for interpretation to another, allowing for communication and data exchange while preserving the architecture of the different systems, rather than replacing it.

In healthcare information systems, it is important to consider preserving the architecture, authorizing access, protecting, and securing the data, maintaining meaning for the correct interpretation of the information by healthcare professionals, guaranteeing focus on relevant patient information, etc. For this reason, the interoperability framework resolves and defines the process followed by all these health information exchange needs.

Therefore, in summary we could say that the successful exchange of information between systems is interoperability, and the transformation of data into information and the design of how to share this information is achieved with interoperability standards.

The interoperability standards define how healthcare information can be exchanged and enable the operational processes, storing and sharing of information between different systems.

2.2. Common Guidelines for eHealth Harmonisation and Interoperability

Information systems have evolved very fast technologically, and this has caused a lot of diversity in their management and architecture. It seems that by defining guidelines (common rules that allow us all to interoperate at least with a common criterion and methodology) by design, despite using different standards (common language and a common set of expectations that enable interoperability between systems and/or devices) according to each solution; can greatly improve this situation in the future and facilitate integration between systems.

The refined eHealth European Interoperability Framework (ReEIF)[1] is positioned as an operational tool kit for implementers and purchasers to deploy eHealth systems. It is intended to be used as a reference guide in calls for proposals and tenders for the Connecting Europe Facility (CEF)[2] deployment, but possibly also for deployment at the national and regional levels. The vision is that the eHealth EIF will be leveraged by the eHealth Network for eHealth deployment that takes place in Member States. The high-level concepts are its governance, principles, agreements, interoperability levels, and high-level use cases.

Interoperability involves many different aspects that must be considered. Aspects such as legislation and guidelines, contracts and agreements between exchanging parties, governance and maintenance, shareable workflows, standardised data elements, semantic and syntactic choices, applications, technical infrastructure, and safety and privacy issues all play a part. Only when all these aspects have been considered, and when all stakeholders are involved in the process, implementation can be successful.

To achieve interoperability between systems, the ReEIF[1] define six layers, each of which addresses key aspects of effective communication and focuses on different standards. This structure is like a hierarchical structure.

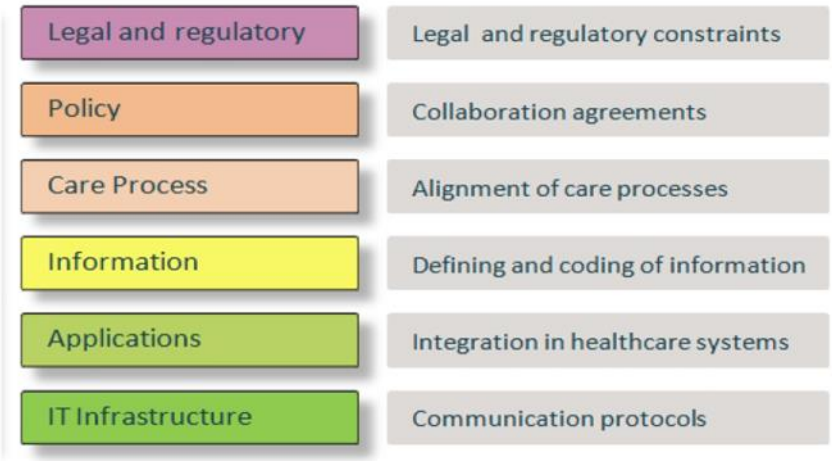


Figure 1 Refined eEIF (ReEIF) model.

[1] ReEIF Overview. (n.d.). from https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b56dffdc&appId=PPGMS

[2] CEF Overview. (n.d.). from https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/connecting-europe-facility_en

European or international projects should follow the ReEIF. INCISIVE follows the ReEIF, but an attempt has been made to group these layers into five to group them into the different tasks distributed among the different partners. Here's how these layers worked in INCISIVE:

- Legal interoperability (Legal and regulatory, Policy): ensures that all the actors operate under the same legal framework, policies, security, and strategies.
- Organizational interoperability (Care Process): ensures that all the actors and transactions of different use cases are standardized in a common and global definition of these scenarios.
- Semantic interoperability (Information): encodes data to maintain the meaning of all information exchanged.
- Syntactic interoperability (Information, Applications): defines how data exchange is to be carried out and definition of its structure.
- Technical interoperability (Applications, IT Infrastructure): defines the technical language that will support syntactic interoperability and relates to interconnection, integration, accessibility, and presentation of data, among others.

For each layer of interoperability, there are different types of interoperability standards. Each of these standards combined with each other address all interoperability requirements.

- Of all existing interoperability standards, it is important to consider the context of use to choose an international, national, or local standard as needed. These are the various uses: International standard: a standard adopted by an international standardisation organisation and made available to the public.
- European standard: a standard adopted by a European standardisation organisation and made available to the public.
- National standard: a standard adopted by a national standardisation organisation and made available to the public.

Depending on the interoperability layer that is being defined, different standards will be used. Each layer of interoperability solves a piece of the data exchange process. The legal layer uses legal and security standards. The organizational layer uses standards such as Integrating the Healthcare Enterprise (IHE)[3] to know that use cases are aligned with the standards. The semantic layer uses standards such as SNOMED CT[4] for clinical data and LOINC[5] for laboratory data. The syntactic layer uses standards such as HL7 FHIR[6] for clinical data and DICOM[7] for medical images.

[3] Integrating the Healthcare Enterprise Overview. (n.d.). from https://www.ihe.net/
[4] Systematized Nomenclature of Medicine-Clinical Terms Overview. (n.d.). from https://www.snomed.org/
[5] Logical Observation Identifiers Names and Codes Overview. (n.d.). from https://loinc.org/
[6] Fast Healthcare Interoperability Resources: Standard of Health Level Seven Organization. (n.d.). from https://www.hl7.org/fhir/index.html
[7] Digital Imaging and Communication In Medicine Overview. (n.d.). from https://www.dicomstandard.org/current

Finally, the technical layer uses standards such as JSON[8] or XML[9] format to represent the structure of messages. Therefore, there are different standards to address each layer of interoperability and the type of information to be exchanged will determine the choice.

It is crucial to use healthcare interoperability standards to ensure consistent data sharing between departments, organizations, levels of care or globally.

[8] JavaScript Object Notation Introduction. (n.d.). from https://www.json.org/json-en.html
[9] eXtensible Markup Language Introduction. (n.d.). from https://www.w3.org/XML/

3. Interoperability framework introduction

3.1. What is an Interoperability Framework

In any project where data exchange occurs between information systems, an Interoperability framework should be added.

The Interoperability framework is a document that includes the analysis of the technical and functional prerequisites of a process/environment or a project, the actors and transactions of each use case involving the data use, the justification of the chosen standards, the methodology and design carried out, the interoperability risk analysis, the interoperability maintenance plan, the publication and reuse of the implementation guides and recommendations and future actions. This document provides a complete overview of the standardization actions carried out in a process/environment or a project.

It is important to consider the use that will be made of the data before starting to define an Interoperability framework, as it is not the same whether the data is used at national, European, or international level. And the standards will also be chosen under this consideration.

The functional and technical requirements for the implementation of the data visualization application and the functional and technical implementation requirements of the interoperability platform that supports the entire specified design are outside the scope of the Interoperability framework.

3.2. What is the recommended methodology for the design and implementation of an Interoperability framework

To start writing and designing an interoperability framework it is important to implement the following steps:

1. Needs Analysis: Analyse the functional, health, technical and security needs that may change the interoperability scenario.
2. Selection of data and information: with the different interest groups, the most common in this phase being health professionals, decide which data to use and which information is needed. Also, which features will use this data.
3. Processes & use cases: Elaborate use case diagrams with the actors and transactions, analyse the datasets, know the processes in which the data will be consulted and which preconditions and postconditions must be met.
4. Search for Interoperability standards and Implementation guides: Select and justify the standards for each layer according to the needs described in points 1, 2, 3 above. Before starting the definition of the interoperability framework, it is also necessary to investigate

whether there are similar implementations that solve the same needs with the already chosen standards. You need to decide if it can be completely reused or if you need to add extensions or profiles. If no implementation guides exist for the same care process or use case, it is our responsibility to create an implementation guide for the definition we design, to help in another future initiative, to contribute to the community, and to facilitate a standard common implementation worldwide for this use case.

5. Risk analysis: Analyse what risks to consider during design and mitigate them when they arise.
6. Maintenance plan: explain how the interoperability framework will be maintained over time, analyse and plan how to deal with standards updates, etc.
7. Specification, Design and Planning of Interoperability Phases: Design each layer of interoperability and plan the design and implementation process from the beginning.
8. Recommendations and future actions: Document some recommendations and future standardization actions for the same interoperability framework or for similar ones that aim at reusing it.
9. Implementation Guide Publication: Explain how to follow the implementation guide to adhere to the system. Also explain how to reuse it for other similar use cases.

These are all the phases that needs to be explained on the interoperability framework:

3.2.1. Needs analysis

This section focuses on the needs and goals of the project, which are important and relevant to making decisions about interoperability design and implementation.

It is necessary to analyse the functional, healthcare, technical and security/legal needs that can change the interoperability scenario.

Therefore, all subsequent decisions take these needs into account.

3.2.2. Selection of data and information

It is important to involve healthcare professionals and/or patients to make a useful and successful solution. The end user is the most important actor in any system, considering that an ideal system that is not used is a system that is useless.

In the health sector, it is necessary to have an exhaustive knowledge of the care process that the system wants to support, it is necessary to know what the relevant information is, what is the useful data and what can help, facilitate, and speed up the work of healthcare professionals.

For this reason, it is essential that healthcare professionals indicate these needs, indicate what information they need in the system and what data is important.

For each term it is important to indicate the value type, whether there is a possible values range or/and units, and an example value.

	Template
Term	**Doses**
Value Type	double
Possible Values	5 - 30
Units	fractions
Example	28

Table 1 Example of the information required for each element.

It is also important to define the data sets, healthcare professionals are responsible for defining which are the appropriate data groups and which events of the care process they correspond to.

This section is one of the most important, as the better definition of the requirements, the better the design of the specification and implementation.

The INCISIVE selection of data and information is explained in section 4.2 of this document.

3.2.3. Processes & use cases

This section should only deal with use cases that require standardized data and that need to be analysed on how to design and implement interoperability.

It is important to define the use cases by indicating the actors and transactions, explaining the entire workflow, and drawing the visual diagrams for each.

INCISIVE has different use cases and roles that use or need standardized data. This use cases are explained in section 4.3 of this document.

3.2.4. Search for interoperability standards and implementation guides

After analysing the needs (Section 3.2.1), defining the information and data to be exchanged (Section 3.2.2), and identifying the use cases, actors, and transactions (Section 3.2.3), it is necessary to select and justify which interoperability standards will be used in each interoperability phase.

In INCISIVE and in most healthcare projects it is necessary to work with different types of data, to share and link images, to make use of prospective and retrospective data, there

are different actors involved in the process, anonymisation, pseudonymisation, and the synthetisation of data are often necessary.

At the same time, the data sharing network must also be considered. In the health sector the network can be very wide, many entities involved in patient care for the same process, many systems of origin and destination, needs for real-time consultation, prediction and prevention, complexity of architecture, cross-border legal and security heterogeneity, need for federated or hybrid systems, etc.

Therefore, when defining the Interoperability framework, it is necessary to choose and justify according to all needs, which standards for each layer will be used and why.

Before starting to make the specification, design, and planning of the interoperability phases, it is also necessary to know if similar implementation guides already exist for the same care process.

This search, selection, and justification for INCISIVE can be found in section 4.4 of this document.

3.2.5. Interoperability Risk analysis

When planning the methodology to be followed for the design and determining the effort for each task, it is necessary to consider the risks to be as realistic as possible and defines how will be mitigated if they arise.

The most common interoperability risks when designing an interoperability framework are:

- Technical or functional requirements may change.
- New needs or new terms to be coded may arise.
- Technologies or architecture development may limit the syntactic interoperability design at some point in the process.
- The versions or dependencies related to the releases of the chosen standards may become obsolete and need an upgrade.
- Changes in the interoperability technical team.
- Lack of documentation about what has been decided and designed about interoperability phases and what still needs to be done.
- Analysis of new recommendations and learning curve.
- Lack of communication with developers while implementing the interoperability specification.
- Lack of support and participation in the configuration, authentication, and security of the servers of each standard.
- Lack of Healthcare professionals support and participation in the selection of data and information.

The interoperability risks that have been mitigated at INCISIVE are explained in section 4.5 of this document.

3.2.6. Interoperability Maintenance plan

When designing an interoperability framework, it is not only necessary to consider the project purpose and needs for implementing interoperability, but also how the interoperability framework will be maintained over time.

This is because any system, standard and specification need to be updated and maintained over time, otherwise it will gradually lose its functionality and benefits.

Therefore, ensuring the compatibility and scalability of the interoperability framework by adopting global standards and specifications makes it possible to use it smoothly in the future as well.

However, any standard will be updated, and we must pay attention, analyse and plan how to deal with the updating of standards.

In addition, the project itself is subject to change, and any change may have an impact on the designed interoperability framework, which also needs to be analysed.

Regarding the maintenance of legal interoperability, it is important to consider the validity of the patients' consent and to plan and define the protocols to be followed when there are requests such as deleting the data, extracting them, some technical change in the processing, etc.

Regarding the maintenance of organizational interoperability, it is important to consider the changes that may occur in the use cases or the implementation of new functionalities and new workflows.

Regarding the maintenance of semantic interoperability, it is important to note that semantic standards are updated periodically, and each update may contain new and/or inactive concepts. Also, we need to consider if there are new concepts during implementation and they need to be codified. And, if some changes in the syntactic or technical interoperability layers affect the semantic encoded codes.

Regarding the maintenance of syntactic interoperability, it is important to note that also syntactic standards are updated periodically, and each update may have changed about the structure definition or the implementation and configuration of the standardized server. For clinical data, therefore with standard specification updates, we need to analyse the changes that impact the implementation and develop a strategy for handling these changes that make it simple and fast to apply the latest standard specification in the future. Also, we need to consider if there is any update of the standard version of standardized server and/or update the standard specification to ensure that standardized

server supports this standard specification. For medical images, we need to consider DICOM[10] updates and PACS[11] version updates.

At the technical interoperability, as mentioned before, the standardized server will also have a series of updates, not only because of the update of the standards, but also their own updates to make better application of the corresponding standard. Most standardized server updates are backward compatible, meaning that documents or systems created with the old version will still operate or be used normally after the standardized server has been updated to the new version. The only thing to note is that we need to check if the standardized server supports the version of the standard specification that we are using and that all the dependencies are updated.

The INCISIVE maintenance plan is explained in section 4.6 of this document.

3.2.7. Specification, design and planning of interoperability phases

This section describes how to implement and design each layer of interoperability and what tasks are needed.

The INCISIVE specification, design and planning of interoperability phases is explained in section 4.7 of this document.

3.2.7.1. Legal Interoperability

Legal interoperability ensures that all actors and systems involved follow a common framework of rules and laws for communicating or exchanging data, in the extent applicable to them, respecting the mapped legal requirements and the laws of their jurisdiction, as well as protecting the data privacy of end users. This is particularly important in cross-border data exchange, where different countries may have different data protection laws and standards.

The main reference for legal compliance standard is the EU's General Data Protection Regulation (GDPR)[12], which requires controllers to process personal data only if they have a legal basis (such as consent) and to follow the seven main principles of data processing. In this sense, GDPR has had a beneficial impact on data security, for patients, doctors, and healthcare providers. In addition to GDPR, the EU is implementing additional safeguards

[10] Digital Imaging and Communication In Medicine

[11] PACS Overview. Arora D, Mehta Y. Use of picture archiving and communication system for imaging of radiological films in cardiac surgical intensive care unit. J Anaesthesiol Clin Pharmacol. 2014 Jul;30(3):447-8. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4152706/

[12] General Data Protection Regulation Overview. (n.d.). from https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en

and rules relating to exchange and sharing of health data, in the Data Governance Act (DGA)[13] and the proposal for European Health Data Space Regulation (EHDS)[14].

While protecting the data privacy of end users, healthcare data security is also needed to protect this data from unlawful breaches of healthcare information security. Because healthcare information may be directly related to a person's private life and health, data minimization techniques, such as use of pseudonymised data or anonymised data may be deployed.

Furthermore, for the protection of the security of healthcare information, it is necessary to adopt security standards to set clear requirements for information security management systems. For example, ISO 27001[15] is a standard for information security management systems that defines information security risks through appropriate measures and controls, while ISO 27799[16] provides best practices for handling data flows in healthcare.

ISO 27799[16] applies to health information (medical images, word and figures, videos) regardless of which systems and methods are used to store and regardless of which communication technologies and protocols are employed to transmit it. Since healthcare sector poses very peculiar information security requirements, ISO 27799 provides guidelines for the definition, implementation, and management of policies and controls for allowing healthcare organizations to assess threats and vulnerabilities and ensure confidentiality, availability, and integrity of health information. It is worth noticing that ISO 27799 does not make recommend any security technology due to the continuous evolution of the technology landscape. This technology neutrality makes ISO 27799 an efficient framework driving and supporting healthcare organizations aiming to protect health information by taking into consideration the specific security risk context.

The National Institute of Standards and Technology (NIST) released Privacy Framework[17] for Improving Privacy through Enterprise Risk Management. This framework aims to provide organizations of all sizes with a common and adaptable approach to privacy risk management that supports the adoption of the “privacy-by design paradigm” while designing and developing systems, applications, and services dealing with personal

[13] European Data Governance Act Overview. (n.d.). from https://www.european-data-governance-act.com/

[14] European Health Data Space Overview. (n.d.). from https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

[15] ISO/IEC 27001 Standard: Information Security Management Systems. (n.d.). from https://www.iso.org/standard/27001

[16] ISO 27799:2016 Standard: Health informatics — Information security management in health using ISO/IEC 27002. (n.d.). from https://www.iso.org/standard/62777.html

[17] Nadeau, E. (2020). Nist privacy framework: a tool for improving privacy through enterprise risk management. https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework_V1.0.pdf

information, the exchange and the communication of information, procedures, and privacy practices, and the cross-organizational workforce collaboration in the context of data processing ecosystem.

3.2.7.2. Organizational Interoperability

Organizational interoperability ensures that the defined process follows a standard process. It is important to analyse the defined workflows and search if similar workflows exist. The goal is to use as much as possible already standardized workflows and use cases to ensure that we are not defining something that already exists or has already been defined and is being used and interpreted around the world. In this way, we streamline the process that can help organizations ensure compliance with industry regulations and standards.

Clinical processes are specific activities, procedures and protocols followed in the healthcare system. The standard clinical process defines the guidelines on which actor and transaction should be considered.

One of the organization layer interoperability standards is Integrating the Healthcare Enterprise (IHE)[18] Profiles which describes how healthcare information systems should share information and provides integrated profiles that describe the management of clinical information, use cases, and specify how to use existing standards such as HL7 FHIR, DICOM, etc. to meet specific clinical processes.

3.2.7.3. Semantic Interoperability

Semantic interoperability is the concept to preserve the meaning of the data to be exchanged, ensuring that they do not lose their meaning and with an unequivocal meaning. In the clinical environment, it is important to understand the type and value of clinical data to create common meaning when people and systems use different words and codes to describe the same thing, because dependent data types and values can generate different medical concepts. If the data is inaccurate or misunderstood, it can cause medical errors. Therefore, there is a need to standardize terminology to represent the meaning of clinical data and help maintain medical concepts and information.

The standards used in this layer are terminological standards, which allow the standardization of data according to specific clinical concepts and the use of coding systems that allow all members of a healthcare system to understand the meaning of the data. There are different types of terminologies for different types of medical data.

[18] Integrating the Healthcare Enterprise Overview. (n.d.). from https://www.ihe.net/

As an example, for clinical, diagnostic, mammography data, etc. SNOMED CT is used, and for laboratory data LOINC is used.

Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT)[19] is the terminology used to code, retrieve, communicate, and analyse clinical data, enabling healthcare professionals to express clinical data accurately and unambiguously.

Logical Observation Identifiers Names and Codes (LOINC)[20] is the terminology used to identify laboratory data such as laboratory tests, measurements, etc.

3.2.7.4. Syntactic Interoperability

Syntactic interoperability is the concept to preserve the architecture and specify the format of the data to be exchanged. The standards used in this layer are the messaging standards that allow the standardization of the sending, receiving, and processing of data between different systems according to the type of data to be exchanged, they allow all members of a healthcare system to speak the same language.

The standards that can be used at this layer are: HL7 v2 or v3, HL7 FHIR, HL7 CDA, OpenEHR, OMOP, DICOM, etc.

HL7 is a set of standards to facilitate the exchange of clinical data. Fast Healthcare Interoperability Resources (HL7 FHIR)[21] organizes the structure with resources, which describe formats and elements of interchangeable health data. It is designed with an HTTP-based RESTful protocol that allows to interact with resources and perform different operations on data, for example, create new resources, update existing resources, search data using filters, etc.

Digital Imaging and Communication in Medicine (DICOM)[22] is the world-recognized standard for medical image interchange, designed for the handling, storage, printing, and transmission of medical images.

Neuroimaging Informatics Technology Initiative (NIFTI)[23] is a data format for the storage of medical images, which can be extended to include additional information such as annotations.

Nearly raw raster data (Nrrd) is a library and file format for the representation and processing of n-dimensional raster data.

[19] Systematized Nomenclature of Medicine-Clinical Terms Overview. (n.d.). from https://www.snomed.org/
[20] Logical Observation Identifiers Names and Codes Overview. (n.d.). from https://loinc.org/
[21] Fast Healthcare Interoperability Resources: Standard of Health Level Seven Organization. (n.d.). from https://www.hl7.org/fhir/index.html
[22] Digital Imaging and Communication In Medicine Overview. (n.d.). from https://www.dicomstandard.org/current
[23] NIFTI Overview. (n.d.). from https://nifti.nimh.nih.gov/

The Segmentation Information Object Definition (IOD)[24] specifies a multi-frame image representing a classification of pixels in one or more referenced images.

From HL7 the HL7 CDA[25] standard is used to produce clinical documents.

Before starting to think about how to define the messages, it's important to search for already existing solutions with the standards selected.

3.2.7.5. Technical Interoperability

Technical interoperability is the concept that defines the technical language that will be used to implement syntactic interoperability and relates to data exchange and storage technologies and protocols.

The implementation of syntactic interoperability is the technical interoperability that facilitates storing, retrieving, querying, and analysing data.

In this layer we decide the technology to use for data exchange such as global programming languages, data exchange protocols, standard specialized servers, standard syntaxes, web services, APIs, etc.

It defines how to store data and how to configure the standardized servers and it also defines the process of using this data and how to query it, as well as the generation of different clinical documents according to the process.

3.2.8. Implementation guide publication

If we follow the entire methodology and go through the phases explained in the previous sections, but do not do this last task of publishing and designing the implementation guide and sharing it with the community of the standards used, we will never be able to generate a standard on similar use cases or care process.

3.2.9. Recommendations and future actions

Finally, it is also important to reflect internally on possible enhancements and future actions to improve the lifecycle of the interoperability platform, the defined implementation guides, and the chosen standards.

This section would be like a letter to the future, either for when a research project goes into production or when an existing platform needs to be updated and maintained following these recommendations.

[24] Segmentation IOD. (n.d.). from https://dicom.nema.org/dicom/2013/output/chtml/part03/sect_A.51.html
[25] HL7 CDA Overview. (n.d.). from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

3.3. Differences between Interoperability framework and Implementation guide

In any use case or process where data is exchanged between information systems, an Interoperability framework should be added. Any Interoperability framework should have one or more implementation guides.

As we have already seen in the previous sections, an Interoperability framework is a document that explains and justifies the process to follow, the design planned, and the decisions made. It is a document that can be fulfilled during the whole lifecycle of a project, and is a kind of history about what needs to be done and what is done, or discarded, etc. This is a mandatory or recommended internal document.

However, an Implementation Guide can be a document, a standard resource, a web guide, etc. The focus is not on how some system was standardized, is more on what is needed to be implemented on a system to be able to exchange data with this one. It is more about developing implementation and knowing how to send the data or how to receive it. This is a mandatory or recommended public document or guide.

It is important not to confuse these two resources to use them correctly and to redirect each stakeholder to the most appropriate document.

4. INCISIVE Interoperability framework

As explained in the section above the Interoperability framework is an internal document within a project or use case. This section only provides an overview of the INCISIVE Interoperability Framework.

Following the methodology explained in section 3.2 these are the tasks done in each phase.

4.1. INCISIVE Needs analysis

The main objectives of INCISIVE are:

- Establish a federated repository of medical images and clinical data that allows secure donation and exchange of data in compliance with ethical, legal and privacy requirements.
- Create an AI-based toolbox for incorporating AI features that improve cancer imaging and enable effective decision-making.
- Enable data providers to have complete control over their data through the integrated federated data storage model, ensuring their autonomy while contributing to the platform.
- Involve stakeholders from different backgrounds including healthcare professionals, patients, policy makers, AI experts, etc.
- **Implement an interoperable data integration service, based on worldwide interoperability standards for medical image and clinical data exchange.**
- Anonymisation of data for privacy protection and compliance with the highest data privacy and security standards.
- Development of AI models combined with a prospective and retrospective dataset, enabling multimodal exploration.
- Enhance explicability by building a Machine Learning based automatic annotation system to produce data for the training of algorithms in AI research.

The project brings together multiple users from around the world who have the potential to benefit in multiple ways. The target users & beneficiaries are:

Data providers benefit from feeding AI models with data to better predict cancer in the future and provide their healthcare professionals with a diagnostic support tool. Oncologists benefit from using AI inference models as a supporting tool for cancer prediction and prevention, and thus provide more effective healthcare to patients by uploading clinical data and medical images of each patient and analysing the results.

Radiologists benefit from the analysis of medical images using AI to detect any abnormalities and improve diagnostic accuracy. With the AI toolbox, routine tasks such as image segmentation and analysis can be automated, reducing the workload of radiologists giving them support.

AI researchers benefit from testing AI models and improving explainability. The federated repository gives them access to a large and diverse dataset of clinical information and medical images, resulting in more diverse and representative training data, improving the performance and accuracy of their AI models, and making it possible to develop more powerful and effective algorithms that advance the research.

Finally, the most benefited is the Patient with better healthcare resulting in more complete information enabling more accurate cancer prediction, surveillance, and follow-up treatment.

Important and relevant needs to making decisions about interoperability design and implementation are:

- Functional and healthcare needs:
  - Prospective and retrospective data
  - Medical images and images annotation
  - Clinical data
  - English language
- Technical and security needs:
  - Federated model / Hybrid model
  - Anonymised data
  - Clinical data and medical image search
  - Using worldwide standards and technology

4.2. INCISIVE Selection of data and information

To define the protocol for the clinical metadata collection and overcome the homogenization challenges in functional, semantics and privacy levels, an iterative procedure took place following the steps:

I. Identification: Proposition of a template per cancer type based on bibliography and medical experience.
II. Review: The templates were circulated through the Data providers, reviewed, and discussed.
III. Merge: A consensus of each template was extracted and discussed in a meeting to resolve homogenization issues.
IV. Redefine: The data providers were asked to provide a sample case. The cases were reviewed for integrity and privacy issues.
V. Standardize: Standardization of the fields content and adopted terminologies.
VI. Review and Refine: The templates were circulated again for verification.

This procedure is described in Figure 2:

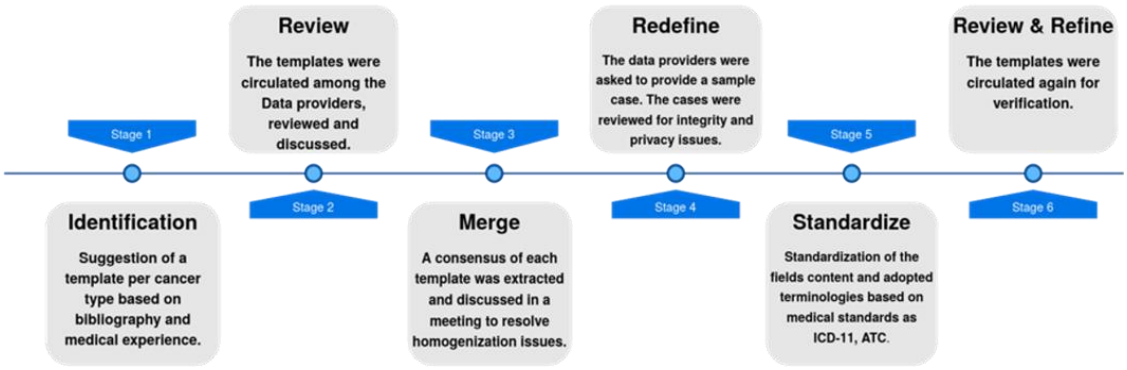


Figure 2 Data collection procedure.

The resulting templates constitute the data collection protocol for the non-imaging data and the basis for the INCISIVE data model. The structure is depicted in brief in the following image. Besides the structure of the templates, and the list of fields of clinical metadata that need to be provided, this procedure defined:

1. The value ranges and allowable types of each field, based on the terminologies.
2. A set of fields that were considered as mandatory. An example of the list of mandatory fields for breast cancer is:
   - **General info**: Age, Gender, Current state.
   - **Baseline**: Image modality (ies), BIRADS classification, TNM staging, Max tumour diameter, BREAST, LOCATION, Pathological lymph-nodes, other breast lesions.
   - **Timepoints**: Image modality (ies), BIRADS classification, TNM staging, Max tumour diameter, BREAST, LOCATION, Pathological lymph-nodes, other breast lesions, Response to treatment.
   - **Treatment**: Type of Treatment.
   - **Histology**: Cancer Type, Grade.
   - **In all above**: Dates (months from diagnosis) & Labels.

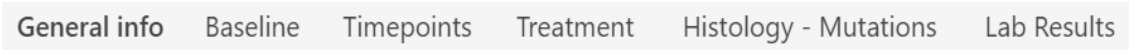


Figure 3 Capture of template tabs.

3. The definition of Timepoints as:
   - **Diagnosis**: will contain all initial exams and information collected during the diagnosis of the pathology. (Month: 0)
   - **After 1st treatment**: will contain exams and information collected during the assessment of the first cycle of treatment. First cycle of treatment is a surgery or a therapy/ combination of therapies that the clinician has considered as the best approach for the patient. In case there are multiple examinations, we will allow for

multiple entries under the tag 1st treatment, but they will be characterized and distinguished by an additional field. (Month: 1-3 approximately, depending on the applied intervention)

- **1st Follow Up**: will contain exams and information collected in two cases: (a) the patient is in remission, and this is a routine follow up, or (b) the patient is in relapse; the therapeutic procedure continues and this timepoint corresponds to a second cycle of treatment. (Month: 4-7 approximately, depending on the applied intervention or the cancer/country specific pathway, guidelines)
- **2nd Follow Up**: will contain exams and information collected in two cases: (a) the patient is in remission, and this is a typical follow up or (b) the patient is in relapse; the therapeutic procedure continues and this timepoint correspond to a third cycle of treatment. (Month: 9-12 approximately, depending on the applied intervention or the cancer/country specific pathway)

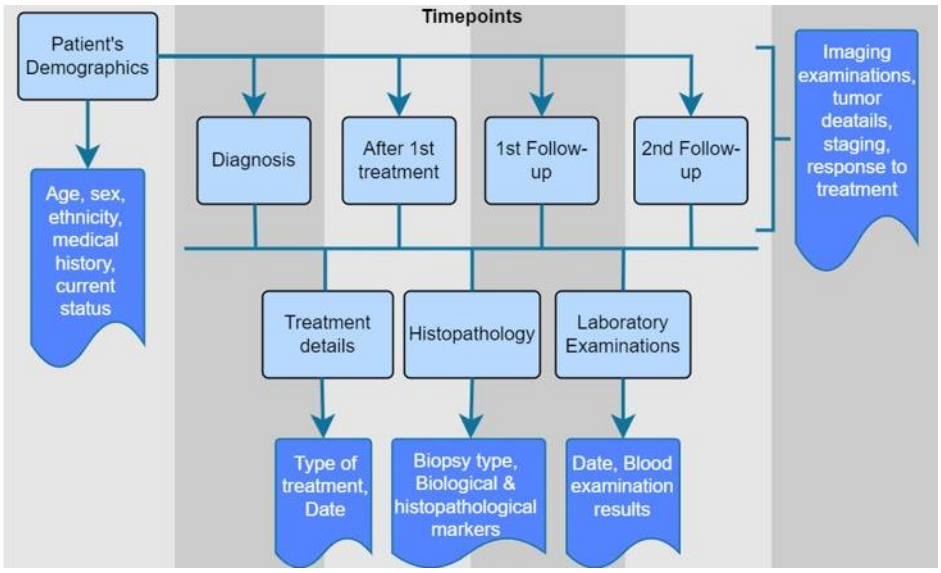


Figure 4 Definition of Timepoints.

An example of the template for breast cancer Baseline tab is:

Patient Number*	Age at diagnosis	Gender	Ethnicity	Birth(s)	pre/post menopause
		0: male 1: Female	1White, 2:Roma, 3:Any other white	No of Births	0:pre, 1: post

Figure 5 Example of template for breast cancer.

The templates that were created for the first phase of data collection (retrospective) were revised and adjusted to the needs of clinicians and developers for the prospective studies. All changes

made were logged and mapped so the initial templates could be converted to the new version without inconsistencies and loss of information.

This conversion took place after the uploading of the prospective data in the central infrastructure. A script was developed that merges the first with the second version of the templates and converts all the values for the fields that changed. The conversion of the templates took place in the central infrastructure and then the data was migrated to the federated nodes.

According to the data collection process, the INCISIVE clinical data is provided by Excel templates defined by Data provider, one for each cancer (Breast cancer, Lung cancer, Prostate cancer, and Colorectal cancer). In this way, the Data provider can insert the mandatory data into the predefined template and upload it. It means that the Data provider has an Excel template with all the input fields to be filled and are shown in section 4.3.

With reference to Excel templates defined by the Data provider, INCISIVE also defined the specification template based on the Excel template to construe HL7 message[26]. In this specification template, it introduced the series of terms that provide not only the name of an input field, but also the possible values of the terms, which is the form of predefined value or free text value or measured value.

The terms come from various phases of healthcare process: For this reason, the classification of the terms is by type of data, such as general information of patient, diagnosis, laboratory results, etc. These classifications are performed in different tags in templates.

- General information: contain terms about demographic information and medical history of patient, also with terms about metadata of message.
- Baseline: contain terms about baseline imaging and TNM staging information.
- Timepoints: contain terms about timepoints imaging information.
- Treatment: contain terms about treatment information of patient during treatment process.
- Histology-Mutations: contain terms about histopathology image details of patient.
- Laboratory results: contain terms about laboratory results of patient.

These are the groups of data defined for the templates:

Cancer Type	Sheet	Data set
Breast, Colorectal and Lung	General info	Demographics
		Medical history
		Metadata

[26] Used HL7 FHIR Standard

	Baseline	Baseline imaging
		TNM staging
	Timepoints	Timepoints Imaging
	Treatment	Treatment Information
	Histology - Mutations	Tumour profile
		Histopathology Image details
	Lab Results	Laboratory results
Prostate	General info	Demographics
		Medical history
		Diagnosis
		Metadata
	Baseline	Baseline imaging
		TNM staging
	Timepoints	Post-operative control
		Recurrence after treatment
	Treatment	Treatment Information
	Histology - Mutations	Biopsy Information
		Final Pathology
		Histopathology Image details
	Lab Results	Laboratory results

Table 2 Groups of data defined for the templates.

*DataSets in blue cell background are not shared by the four cancer types.

It is very important to take these datasets into account in semantic and syntactic interoperability phases to know the limitations that will arise throughout the process in the specification of the elements for each term.

4.3. INCISIVE Processes & use cases

INCISIVE has different use cases and roles that use or need standardized data. This section only shows use cases that require standardized data and that need to be analysed on how to design and implement interoperability. The rest of the use cases can be consulted in INCISIVE Deliverable 2.3 User Requirements Definition and System Design.

These use cases are the ones that use or need standardized data:

- Uploading data as a Data Provider

- Querying data as an AI Researcher
- Deploying an AI Engine as an AI Researcher
- Uploading data for AI Service as a Healthcare Professional
- Download clinical report of an AI Service as a Healthcare Professional

For Model training:

- Use case 1: Uploading data as a Data Provider

  The data provider uploads the data to its node. Data providers can choose two solutions, install a federated node, and manage the data locally, or use the centralized node offered by INCISIVE that acts as another federated node. Data providers are responsible for using the quality check tool and image anonymisation tool before uploading data to the node. The data includes medical images and clinical data, each of which has its own repository and workflow. All this data is processed inside the node.

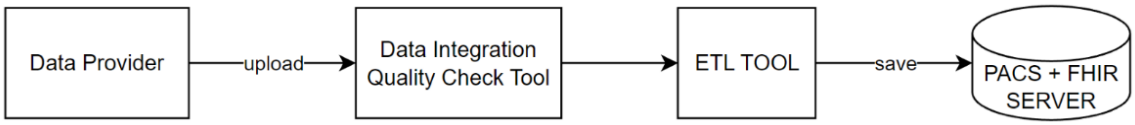


Figure 6 Uploading data as a Data Provider workflow.

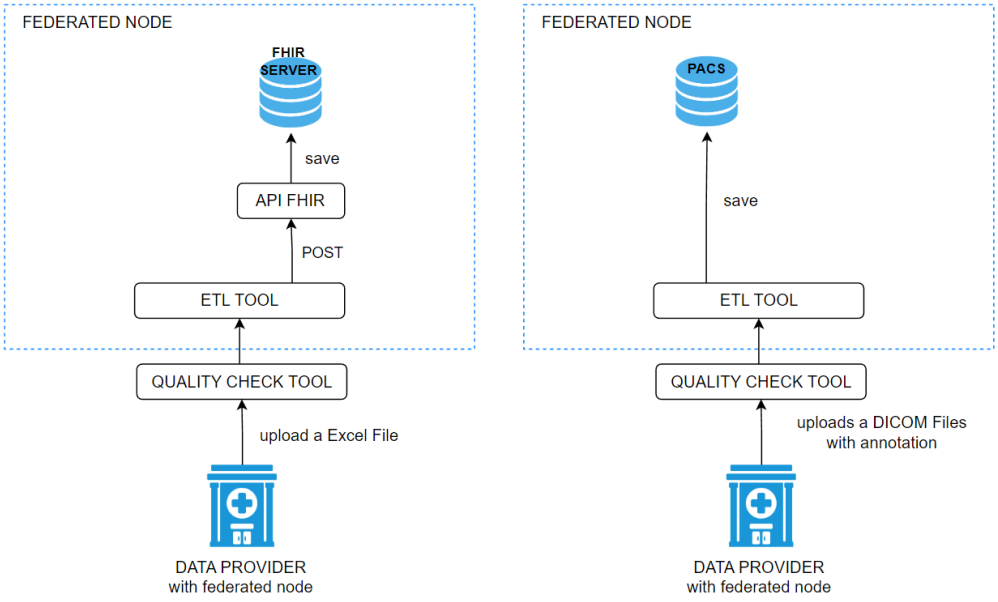


Figure 7 Uploading data as a Data Provider case use.

The diagram above shows details of uploading data in federated node. It's the same process for clinical data as for medical images, but the folder structure is different.

Clinical Data Uploading:

ACTORS	• Data provider • Quality check tool • Federated Node or Central Node as Federated Node (ETL Tool, API FHIR, FHIR Server)
TRANSACTIONS	• The Data provider use the quality check tool locally. • The Data provider set and install the Federated Node or use the Central Node as a Federated Node. • The Data provider uploads the Excel file with data of all the patients, that have given their consent, to the ETL Tool. • The ETL Tool transforms this data into HL7 FHIR message and perform a POST to the FN FHIR Server. • The FHIR server saves this data to be queried locally later by the AI engines.

Table 3 Clinical Data Uploading use case.

INCISIVE provides Data providers with an Excel template with a header indicating the terms to fill in with their patient data. Below is an example of Excel that has been uploaded by a Data provider with values from different patients: Each row corresponds to a patient, identified by a patient id. As the data are of different types it is the responsibility of the Data provider to use the quality check tool to detect possible errors before uploading the file to the Federated Node.

Patient Number	Exam Label	Date	Leukocytes	Hb	HCT	PLT	Blood sugar	Insulin	Urea	Creatinine	Uric acid	Potassium	Sodium	Calcium	Cholesterole	Triglycerides	HDL	LDL
P1HCS	0	1	5,9	12,9	39,2	293	98	n/a	21	0,6	2,5	3,8	135	9,1	141	61	55	73
P1HCS	1	7	6	12,5	39,3	280	82	n/a	22	0,7	3	4	138	9,5	145	78	67	78
P1HCS	2	11	5,8	12,6	39,5	215	84	n/a	20	0,6	3,2	4,1	141	9,2	167	75	71	82
1DISBA	0	0	8,2	10.3	30.8	226	138		49	0.7	4.7	4.7	140	10.2	141	95	N/A	N/A
1DISBA	1	8	8.9	9.5	29	201	166		N/A	1.1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1DISBA	2	14	9.6	9.7	32.5	230	174		N/A	0.93	6.1	4.8	138	9.7	N/A	N/A	N/A	N/A
Br001GOC	0	N/A	27,9	13	41	194	N/A	N/A	50	0,81	N/A	5,6	143	N/A	N/A	N/A	N/A	N/A

Figure 8 Example of Excel template uploaded by Data provider.

Image Data Uploading:

ACTORS	• Data provider • Quality check tool and anonymisation image tool • Federated Node or Central Node as FN (ETL Tool, API FHIR, FHIR Server)
TRANSACTIONS	• The Data provider use the quality check tool and anonymisation image tool locally. • The Data provider set and install the Federated Node or use the Central Node as a Federated Node. • The Data provider uploads the DICOM files with images of all the patients, that have given their consent, to the ETL Tool. • The ETL tool saves these images and annotations to the PACS to be queried locally later by the AI engines.

Table 4 Image Data Uploading use case.

- Use case 2: Querying data as an AI Researcher

  The AI Researcher can query data to train AI models by searching based on various filters in Central node's INCISIVE UI. When this search is done, the Central node queries within each Federated node and then returns the results of all local searches to the Central node to be viewed by the AI Researcher.

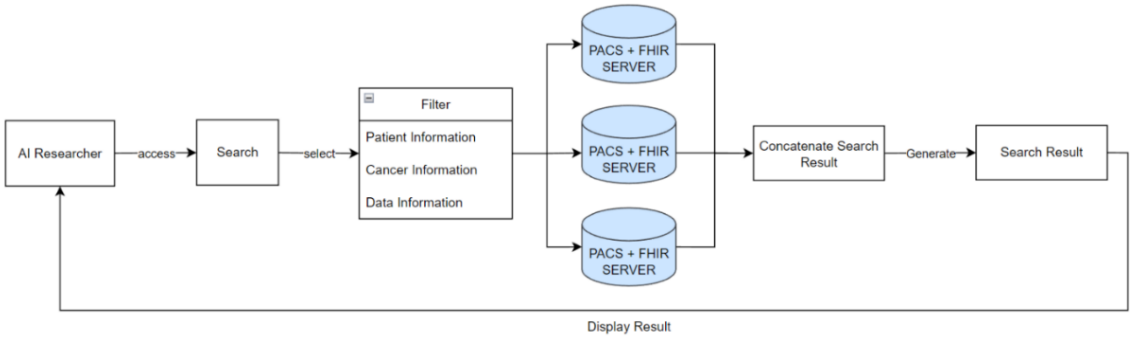


Figure 9 Querying data as an AI Researcher workflow.

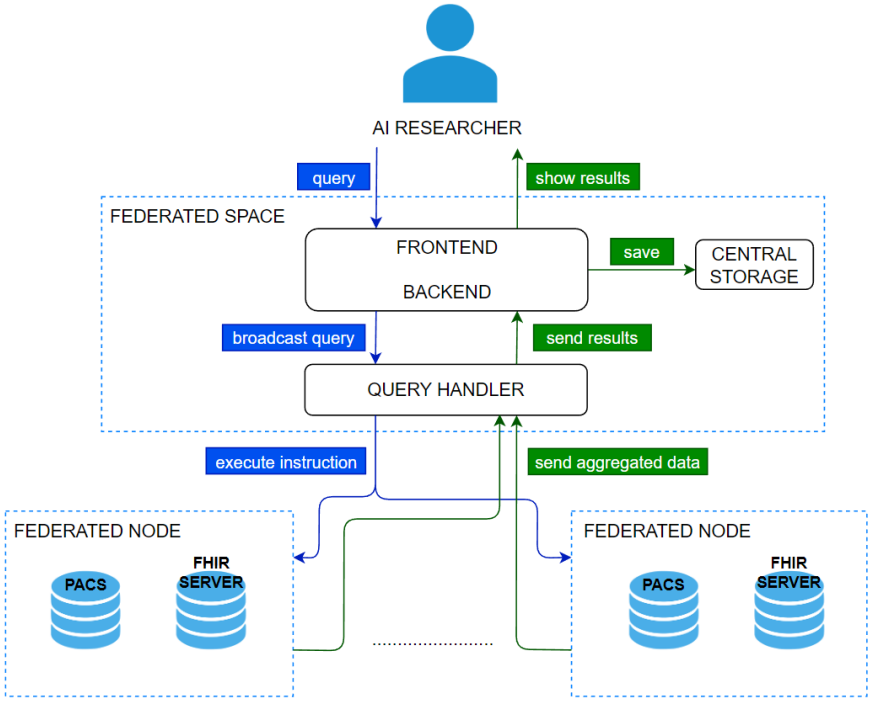


Figure 10 Querying data as an AI Researcher case use.

The diagram above shows details of the query process at a technical level.

ACTORS	• AI Researcher

	• Frontend • Backend • Query Handler • Federated node • Central Storage
TRANSACTIONS	• AI Researcher query with different filters on the INCISIVE frontend. • The Backend broadcasts the query of the frontend and sends to the Query handler. • The Query Handler executes the instruction to all available Federated Nodes. • Each Federated Node send aggregated data of all their patients to the Query Handler. • The Query Handler sends the results to the Frontend. • The Frontend shows a summary to the AI researcher and saves the query to Central Storage.

Table 5 Querying data as an AI Researcher.

- Use case 3: Deploying an AI Engine as an AI Researcher

  With previous search results, an AI Researcher can use this data to train the model.

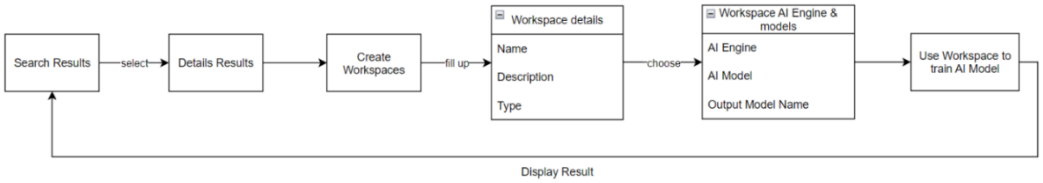


Figure 11 Deploying an AI Engine as an AI Researcher workflow.

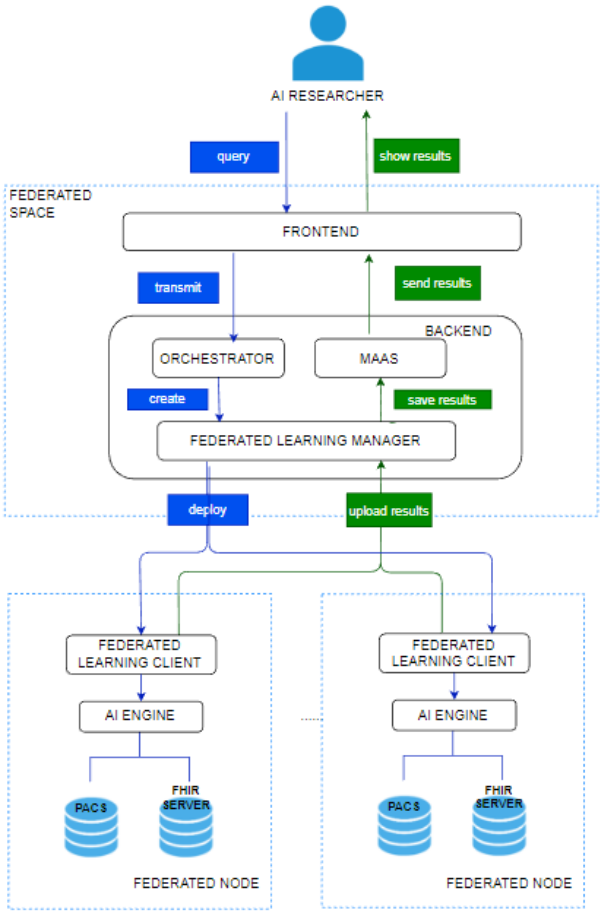


Figure 12 Deploying an AI Engine as an AI Researcher case use.

ACTORS	• AI Researcher • Frontend • Orchestrator • MAAS • Federated learning manager • Federated learning client • AI Engine • PACS • FHIR Server

TRANSACTIONS	• AI Researcher selects previous performed data query on the Frontend. • The Frontend transmits the query to the Orchestrator, and this create the Federated learning manager. • The Federated Learning manager deploy to each federated learning client, and this deploys the AI Engine in the Federated Nodes. • The AI Engine obtains the result of the Model Training that used the required data from PACS and FHIR Servers. The AI Engine uploads this result to the Federated learning client, which then sends it to the federated learning manager. • The Federated learning manager save the result to the MAAS and send it to the Frontend. • The Frontend shows the results.

Table 6 Deploying an AI Engine as an AI Researcher.

For Model inference:

- Use case 1: Uploading data for AI Service as a Healthcare Professional

  Healthcare professionals can use different AI services on the INCISIVE platform to make Model Inference and obtain a result and a clinical report for a specific patient. Healthcare professionals upload an Excel with some clinical data of a patient and some DICOM images of the same patient to the INCISIVE UI. This data is fed into the AI models which returns the results and the patient report.

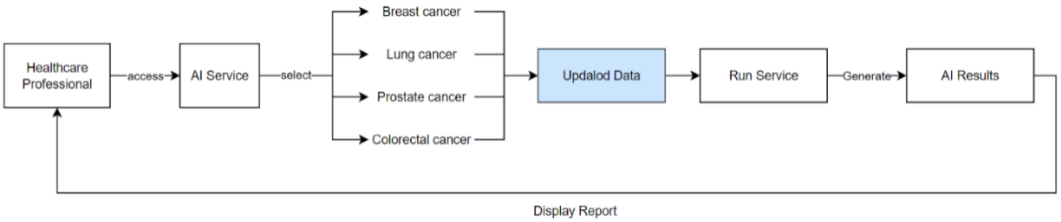


Figure 13 Uploading data for AI Service as a Healthcare Professional workflow.

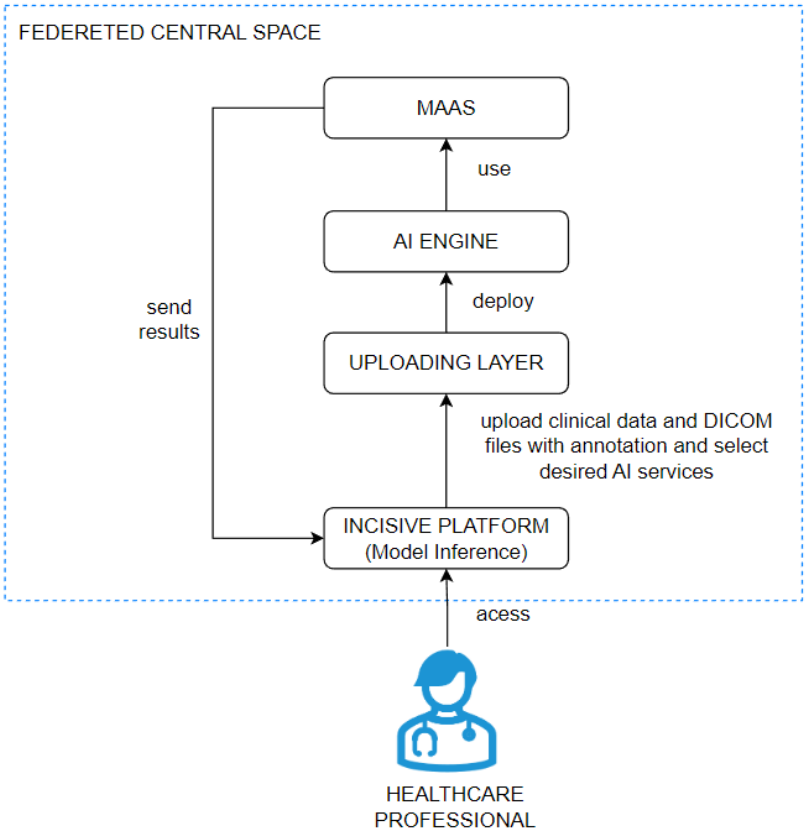


Figure 14 Uploading data for AI Service as a Healthcare Professional case use.

ACTORS	• Oncologist (Healthcare professional) • INCISIVE Platform • Uploading layer • AI Engine • MAAS
TRANSACTIONS	• The Oncologist uploads the clinical data and DICOM files with annotation that wants to use for performing inference of an AI model in the INCISIVE frontend. • The backend receives data to the Uploading layer. • The Uploading layer deploys AI Engine to perform inference using MAAS. • The MAAS generates the results of the inference and sends them to the Frontend. • The Frontend show the results of the inference.

Table 7 Uploading data for AI Service as a Healthcare Professional.

- Use case 2: Download clinical report of an AI Service as a Healthcare Professional

With the previous results, the healthcare professional can download the clinical report and choose which clinical report format they want to download, either as an XML file or as a PDF file or both. The XML file is standardized using HL7 CDA, which means the provider can process this clinical report in their health information system. The PDF file, on the other hand, only allows for viewing on their device.

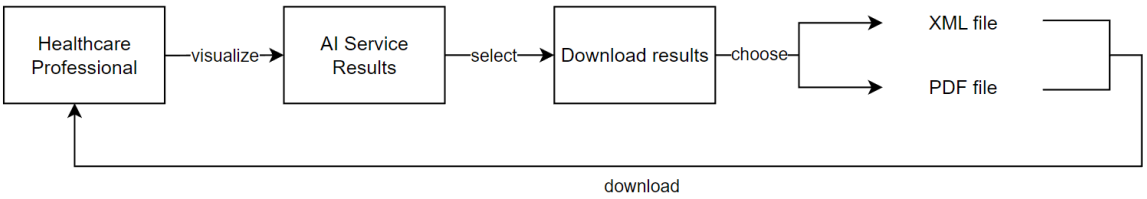


Figure 15 Download clinical report of an AI Service as a Healthcare Professional workflow.

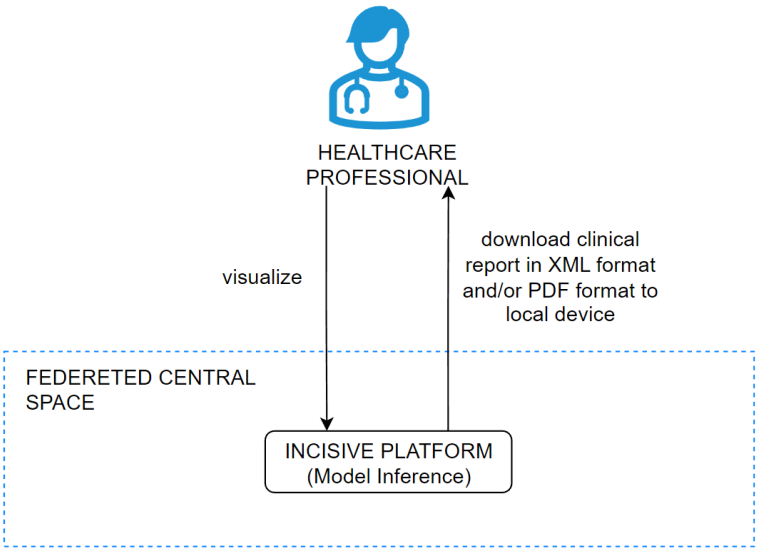


Figure 16 Download clinical report of an AI Service as a Healthcare Professional.

ACTORS	• Oncologist (Healthcare professional) • INCISIVE Platform
TRANSACTION	• The Oncologist downloads clinical reports in XML and/or PDF format on the INCISIVE platform. • The downloaded clinical report is saved on their local device.

Table 8 Download clinical report of an AI Service as a Healthcare Professional.

4.4. INCISIVE Search for interoperability standards and implementation guides

Given the nature of the project, obligatory privacy and ethics requirements needed to be considered at every stage of the lifecycle of INCISIVE, in all activities related to collecting, storing, sharing, and using health data. The identification of the legal requirements and their translation into actionable recommendations was performed during WP7 work. The outcome of the selection of the relevant legal acts was presented in Deliverable 7.3 Data Donation Legal Framework.

Based on the needs we have discussed about the use cases defined in INCISIVE, the IHE has some domains related to radiology and oncology, and some technical workflows regarding the use of results for the AI For this reason, it seems that INCISIVE should consider or analyse whether it can follow the IHE standard.

Based on the INCISIVE needs, SNOMED CT and LOINC were chosen for their wide application worldwide and the possibility to code most of the INCISIVE clinical data. These two standards that we have mentioned in section 3.2.7.3 are supported by their international organizations that maintain them. Therefore, if we cannot codify some of the terms, we can try to request the creation of a concept or use different techniques, for example, post-coordination in SNOMED CT standards to express the required medical concepts by combining multiple concepts.

As INCISIVE is a federated system it needs to make sets of clinical data queries, for this reason Fast Healthcare Interoperability Resources (HL7 FHIR) seems a good choice considering that it organizes the structure with resources, which describe formats and elements of interchangeable health data and which is designed with an HTTP-based RESTful protocol that allows you to interact with resources and perform different operations on data, for example, create new resources, update existing resources, search data using filters, etc.

As INCISIVE uses medical images that need to be stored, exchanged and consulted, DICOM has been chosen.

In INCISIVE, NIFTI has been chosen for medical image annotations.

HL7 CDA standard is also used, as INCISIVE needs to produce a clinical document resulting from AI services for model inference, a pdf template is used to be visualized in the UI[27] and an HL7 CDA to be exported and used by data providers.

In summary, in INCISIVE project, to achieve worldwide exchangeable data, it has been decided to use HL7 FHIR for clinical data, DICOM for medical images and NIFTI for medical image annotation.

FHIR Server was chosen for clinical data considering that AI Engines need to query this data quickly and in a standardized way, and PACS was chosen for medical image and annotations to process

[27] User Interface

the DICOM files; the use of the RESTful methodology will make it much easier for AI Models and AI Services to query this data.

At INCISIVE, the IBM FHIR Server[28] has been chosen to store clinical data, as it is an open-source Java solution that supports the processing, validation, and storage of healthcare data according to the HL7 FHIR specification. It was considered when making the decision, that the ETL Tool is written in Java and that the IBM company is a global company to provide maintenance support.

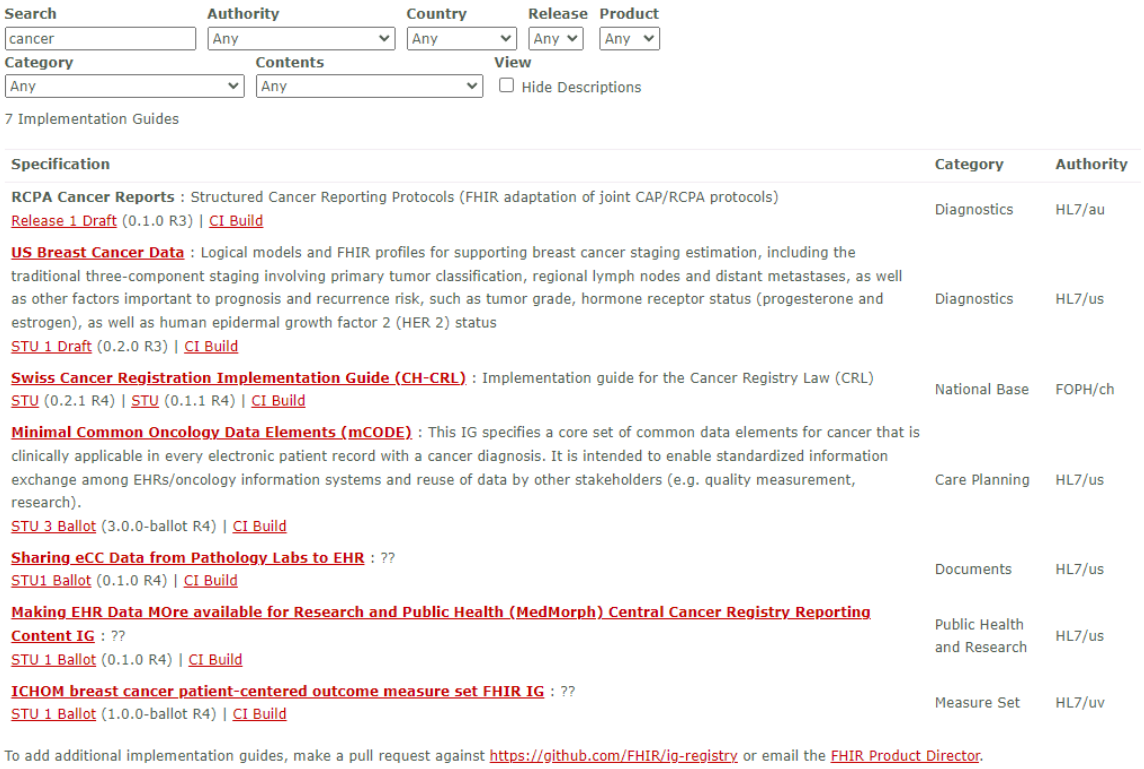
At INCISIVE, the Orthanc PACS[29] was chosen to store the medical images, as it is a free, open source, lightweight DICOM server for medical images. It has been considered that its architecture is light and self-contained, which means that no complex database administration is required, nor the installation of third-party dependencies, and because it provides a RESTful API.

Implementation guides were searched for existing implementations and possible starting points to define use cases for this project. INCISIVE searched for cancer and AI prediction solutions implemented with the standards selected. The different organizations that maintain these standards provides a repository of implementation guides for integrated systems that not only ensure that existing systems can communicate effectively and reduce the risk of errors, but also facilitate the integration of new systems in the future by providing clear instructions on how to integrate systems and which data format structures should be used.

It has also helped us draw the conceptual map of resources, terminology, profiles, and extensions.

On the official page of the FHIR Foundation there is a section Registry of Implementation Guides[30] where you can find all the information related to the registry of existing implementation guides. INCISIVE searched for "cancer" and found 7 implementation guides related:

[28] IBM FHIR Server. (n.d.). from https://www.ibm.com/downloads/cas/EWNW6QE3?mhsrc=ibmsearch_a&mhq=FHIR%20Server
[29] Orthanc - DICOM Server. (n.d.). from https://www.orthanc-server.com/
[30] Implementation Guide Registry. (n.d.). from http://fhir.org/guides/registry/



Search: cancer | Authority: Any | Country: Any | Release: Any | Product: Any
Category: Any | Contents: Any | View: Hide Descriptions

7 Implementation Guides

Specification	Category	Authority
RCPA Cancer Reports : Structured Cancer Reporting Protocols (FHIR adaptation of joint CAP/RCPA protocols) Release 1 Draft (0.1.0 R3) | CI Build	Diagnostics	HL7/au
US Breast Cancer Data : Logical models and FHIR profiles for supporting breast cancer staging estimation, including the traditional three-component staging involving primary tumor classification, regional lymph nodes and distant metastases, as well as other factors important to prognosis and recurrence risk, such as tumor grade, hormone receptor status (progesterone and estrogen), as well as human epidermal growth factor 2 (HER 2) status STU 1 Draft (0.2.0 R3) | CI Build	Diagnostics	HL7/us
Swiss Cancer Registration Implementation Guide (CH-CRL) : Implementation guide for the Cancer Registry Law (CRL) STU (0.2.1 R4) | STU (0.1.1 R4) | CI Build	National Base	FOPH/ch
Minimal Common Oncology Data Elements (mCODE) : This IG specifies a core set of common data elements for cancer that is clinically applicable in every electronic patient record with a cancer diagnosis. It is intended to enable standardized information exchange among EHRs/oncology information systems and reuse of data by other stakeholders (e.g. quality measurement, research). STU 3 Ballot (3.0.0-ballot R4) | CI Build	Care Planning	HL7/us
Sharing eCC Data from Pathology Labs to EHR : ?? STU1 Ballot (0.1.0 R4) | CI Build	Documents	HL7/us
Making EHR Data MOre available for Research and Public Health (MedMorph) Central Cancer Registry Reporting Content IG : ?? STU 1 Ballot (0.1.0 R4) | CI Build	Public Health and Research	HL7/us
ICHOM breast cancer patient-centered outcome measure set FHIR IG : ?? STU 1 Ballot (1.0.0-ballot R4) | CI Build	Measure Set	HL7/uv

To add additional implementation guides, make a pull request against https://github.com/FHIR/ig-registry or email the FHIR Product Director.

Figure 17 Implementation guides related to cancer of official page of the FHIR Foundation.

The "US Breast Cancer Data"[31] and "mCode"[32] implementation guides have been considered, although they have not been reused given the number of differences in terminology needs and the HL7 FHIR version, as INCISIVE uses the latest stable version of HL7 FHIR which is R4.

It has also been decided to develop an implementation guide for the defined messages and that will be published in this register by making the corresponding request.

4.5. INCISIVE Interoperability Risk Analysis

In INCISIVE Interoperability framework a risk analysis was done to prevent some of the common interoperability risks. However, the following risks have been mitigated:

Risk description	Impact

[31] HL7 FHIR Implementation Guide: Breast Cancer Data, Release 1 - US Realm (Draft for Comment 2). (n.d.). from http://hl7.org/fhir/us/breastcancer/2018Sep/

[32] HL7.FHIR.US.MCODE\Home - FHIR v4.0.1. (n.d.). from https://build.fhir.org/ig/HL7/fhir-mCODE-ig/

The need to link the path of the image with the HL7 FHIR messages, need to differentiate terms of time points, and to send more than one time point per patient. Pass from only federated infrastructure to hybrid infrastructure.	Medium
New terms, deleted terms, modified terms for prospective and retrospective templates.	Low
Transaction Tracker instead of ATNA IHE profile[33].	Low
FHIR version R4 to R4B to R5 analysis.	Medium
Some changes in the interoperability technical team.	Medium
Elaborate the IHE Actor Transaction Diagram and reported it.	Medium
XML to JSON was performed.	Low
Support the configuration and query of the servers and add the terminology resources.	Medium
The need of a pdf template and a HL7 CDA Clinical Report for model inference.	Medium

Table 9 Risk Analysis table.

4.6. INCISIVE Interoperability Maintenance plan

In INCISIVE Interoperability framework a maintenance plan was done about the standards and interoperability platform.

These are some of the important parts to consider:

For terminological codes be aware of what to do if the concept used in INCISIVE, SNOMED CT or LOINC is disabled and update it following the directions of the official standards.

For clinical data, therefore with HL7 FHIR specification updates, analyse the changes that impact on our implementation and develop a strategy for handling these changes that makes it simple and fast to apply the latest specification in the future. Also, consider if there is any update of the version of FHIR server and/or update the FHIR specification to ensure that FHIR server supports this specification and that all the dependencies are updated.

For medical image, DICOM updates and PACS version updates must be considered.

4.7. INCISIVE Specification, design and planning of interoperability phases

This section describes how each layer of interoperability was implemented and designed and what tasks have been done in the INCISIVE project.

[33] ATNA IHE Overview. (n.d.). from https://wiki.ihe.net/index.php/Audit_Trail_and_Node_Authentication

4.7.1. INCISIVE Legal interoperability

In this process, INCISIVE needs to comply with relevant laws, regulations, and standards when processing and sharing medical data, and to adopt appropriate technical measures to ensure data security and privacy protection. Achieving data sharing maximizes the utility and impact of the collected data, makes collaboration between different researchers easier and increases the transparency of research results.

To facilitate the GDPR[34] and legal compliance for the stakeholders, project designed uniform rules for the data providers to participate in the INCISIVE platform data sharing, as well rules for the data users to benefit from the data for research purposes. According to them, each data provider in INCISIVE must examine whether they are able to ensure compliance with GDPR requirements when providing data to INCISIVE. It is the responsibility of the data provider to define the data they wish to share on the INCISIVE platform and verify whether that data requires patient consent or ethical approval. They can only upload data in the format defined by INCISIVE.

They also need to ensure data quality standards which were determined by the project.

The INCISIVE platform keeps the processing of personal data to a minimum and therefore requires the data providers to take anonymisation and pseudonymisation measures. Both the clinical data and medical images, contain a lot of personally identifiable information such as patient names, ID numbers. To minimize the processed data and ensure privacy protection of the patients' privacy de-identification (pseudonymisation or anonymisation) by data provider is required for sharing and using these data, i.e., removing or replacing personally identifiable information, making it reasonably unlikely for patients to be identified. Medical images, on the other hand, use DICOM files, as all DICOM files are modified before being uploaded to the repository, so that no personal information about the patient or doctor is revealed. This data can only be used in the secure environment of the platform and cannot be downloaded or copied to an external location.

For users, the INCISIVE data donation framework uses a registered access model for data protection, e.g. researchers need to request access to the content of the repository. This means that to access the data, users need not only to register on the platform, but also to apply for access rights. The users must accept the use conditions. These rules define detailed use limitations, such as restrictions on re-identification attempts, and put in place acknowledgement requirement for the users in exchange for being able to use the data for their benefit. Moreover, the use of the data will be recorded via blockchain

[34] General Data Protection Regulation Overview. (n.d.). from https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en

mechanisms. This means that to access the data, users need not only to register on the platform, but also to apply for access rights.

The detailed explanation on the rules of GDPR, as well as DGA and EHDS in the context of INCISIVE were provided in Deliverable 7.3 Data donation legal framework.

4.7.1.1. IEC and ISO's International Standards

INCISIVE interoperability components should be in accordance with relevant medical device and software standards to ensure their safe operation. Specifically, the relevant standards are:

- ISO 14971 Medical Devices Risk Management Assessment[35]
- IEC 62304 Medical device software — Software life cycle processes[36]

Other standards, including EN 60601[37] family of norms that normally apply to medical devices, are not directly applicable to INCISIVE as it has no direct physical realization (and hence no inherent mechanical or electrical risks to patients, operators, or other persons. If the INCISIVE platform is to be considered a medical device, it's realisation and operation is also covered by the Medical Devices Regulation (EU) 2017/745 – MDR[38].

The ISO 14971[34] Medical Devices Risk Management Assessment standard sets out the requirements for managing risks associated with the use of medical devices and software used in healthcare. The interoperability contained within/among proposed INCISIVE system components as well as INCISIVE components and other medical devices which integrate with it during the processes of data storage, retrieval and service requesting, give rise to specific risks to patients.

One of these risks in INCISIVE could be that the patient data arrives in the system in an incorrect form and that consequently the AI models return an incorrect diagnosis. These errors leading to miss-identification of patients and their diagnoses can lead to hazardous situations in which patient are assigned incorrect diagnoses and can receive treatment for conditions they do not suffer from or delay in receiving of correct treatment.

[35] ISO 14971:2019 - Medical devices — Application of risk management to medical devices. (n.d.). from https://www.iso.org/standard/72704.html

[36] IEC 62304:2006 - Medical device software — Software life cycle processes. (n.d.). from https://www.iso.org/standard/38421.html

[37] UNE-EN 60601-1:2008 Equipos electromédicos. Part 1: Requisito... (n.d.). from https://www.en.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0041083

[38] Medical Device Regulation (MDR) | TÜV SÜD. (n.d.). from https://www.tuvsud.com/en/industries/healthcare-and-medical-devices/medical-devices-and-ivd/medical-device-market-approval-and-certification/medical-device-regulation

To ensure compatibility with the ISO 14971 standard, the INCISIVE software and interoperability components have to be subjected to an integrated risk analysis process that identifies all possible hazards associated with the operation of the system on patient and healthcare data. Initially, acceptability criteria for all possible health risks associated with the use of INCISIVE components must be defined.

Next step is the Risk Analysis process in which all possible hazards must be classified for their severity, magnitude of the effect on patient's or other person's health and well-being. Subsequently, all situations leading to these hazards must be identified and assessed for their probability of occurrence given the current design and known experience and practice with the designed software components. The final step in the risk analysis process is the final assessment of risk acceptability.

Further, in the risk management process that follows on from the risk analysis, all identified risks that exceed acceptability criteria have to be managed by devising minimisation techniques and control measures to ensure they are minimised to acceptable levels and are controlled in order not to increase their probability of occurrence over time.

All previously outlined actions must be documented in the INCISIVE risk file. Given that many of the components will be developed and their development managed by different partners, these risk management aspects could be covered by the risk files of individual system components with a section on the risk management of their interoperability.

IEC 62304 Medical device software — Software life cycle[35] processes defines regulations applicable to the development of software components within INCISIVE and their interoperability. This means documenting the process of requirement specification collation, development planning, development task assignment and realization as well as product functionality verification and validation. All the steps in this process must be documented in the development system and be part of an accessible and auditable technical file for the INCISIVE platform.

The final stages of this process, verification, and validation, imply integrated testing of the produced prototypes by development teams (or their testing components) and intended platform users (doctors and other healthcare professionals) respectively. These processes need to be planned and documented. At their end, development reviews must be conducted to identify potential risks that have not been managed in the development of the platform. Verification and validation tests must specifically cover the interoperability aspects and functions of the INCISIVE platform with other relevant systems.

Furthermore, any changes to the developed platform must be treated as additional development activities and documented in the same manner as the initial development. Records of the development process must be kept in technical files. Given that different

components of the INCISIVE platform will be developed, and their development will be managed by different partners, these technical files would be partner specific.

4.7.2. INCISIVE Organizational Interoperability

INCISIVE wants to be aligned to the IHE global standard, which is why we analysed the IHE workflows and use cases.

IHE creates and operates a process through which interoperability of healthcare IT systems can be improved. IHE is organized in clinical and operational domains. Each domain develops and maintains its own set of Technical Framework documents.

IHE technical frameworks[39] are detailed documents which specify the integration profiles and associated actors (systems) and transactions.

IHE integration profiles[40] describe a clinical information need or workflow scenario and document how to use established standards to accomplish it. A group of systems that implement the same integration profile address the need/scenario in a mutually compatible way.

INCISIVE works in radiology and oncology domains for the use of AI result. The domain of IHE that matches so fair is Radiology, therefore the technical framework that we need to analyse is "IHE Radiology Technical Framework AI Results"[41].

A document on the IHE Actor Transaction Diagram Guide was produced to include work for organizational interoperability. The document describes details of the IHE use cases and scenarios, actors, transactions, and diagram in the IHE Radiology Technical Framework AI Result. Also, it describes the details of the INCISIVE analysis of scenarios and use cases accordance to IHE standards. After the analysis, we concluded that INCISIVE is not that far from IHE standard, as it complies with most transactions (RAD-18, RAD-29, RAD-108, RAD-137, RAD-16, RAD-31, RAD-129, RAD-108) except for two transactions of the IHE standards (RAD-45[42] and RAD-107[43]) that INCISIVE orchestrator or the central node storage does not comply with.

The INCISIVE IHE report was published in INCISIVE Annex 5 in Deliverable 6.2 INCISIVE Integrated Prototypes – Second Version.

[39] Technical Frameworks - IHE International. (n.d.). from https://www.ihe.net/resources/technical_frameworks/
[40] Profiles - IHE International. (n.d.). from https://www.ihe.net/resources/profiles/
[41] Radiology Technical Committee, I. (2023). IHE_RAD_TF_Rev21-0_Vol2_FT_2023-06-15. https://www.ihe.net/uploadedFiles/Documents/Radiology/IHE_RAD_TF_Vol2.pdf
[42] RAD-45 (Retrieve Evidence Docs): A test document user (image viewer, report creator, or report reader) or image document consumer requests and retrieves a test document from the image archive or image document source.
[43] RAD-107 (WADO-RS Retrieve): The transaction accesses DICOM SOP Instances via an HTTP interface.

4.7.3. INCISIVE Semantic Interoperability

In this process, INCISIVE clinical data was semantically encoded in a way that gets everyone within a system to speak the same language and understand the meaning of the data.

Templates contain input fields about clinical elements and laboratory elements, it was distinguished between these two types of data and used SNOMED CT or LOINC according to the medical concepts.

It is very important to have a good definition of the medical concept that must be encoded. For that reason, in this process, we have contacted several times some healthcare professionals to validate our interpretation of the medical concept to be sure that we encode correctly.

The final semantic encoding result is:

	Terms				
	Breast cancer	**Lung cancer**	**Prostate cancer**	**Colorectal cancer**	**Total**
SNOMED CT	154	133	88	111	486
LOINC	38	41	37	40	156
N/A	3	2	1	2	8
Total	195	176	126	153	**650**

Table 10 Semantic encoding result count.

We encoded 486 terms using SNOMED CT, 156 terms using LOINC, and 8 terms were standardized directly by HL7 FHIR message.

All the SNOMED CT and LOINC term codes and their description used by INCISIVE are in ANNEX 1. This coding may be useful for other projects using the same terms. All templates and encoded terms are also in INCISIVE Interoperability framework.

4.7.4. INCISIVE Syntactic Interoperability

With DICOM and HL7 FHIR, INCISIVE achieves a Common Data Model (CDM)[44] that allows receiving, storing, and processing information from multiple sources, multiple data providers, as a one standard way.

[44] CDM Overview. Weeks, J., & Pardee, R. Learning to Share Health Care Data: A Brief Timeline of Influential Common Data Models and Distributed Health Data Networks in U.S. Health Care Research. EGEMs. 2019 Mar; 7(1), 4. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6437693/

4.7.4.1. DICOM and annotation standardization process

DICOM[45] standard has been chosen, for medical images, which is the global standard for medical image exchange, and the NIFTI[46] standard for annotation images.

These are the steps followed to prepare the Medical Images for the AI Models:

- Anonymisation of DICOM Images
- Annotate the DICOM Images
- Transform Annotation DICOM to NIFTI (The details of this step were presented in deliverable DICOM standard for image Annotations Description and document Standards for Image Annotation)

There are differences in the definition and standardization process between medical images and clinical data. Clinical data is uploaded already anonymised, but medical images must go through the INCISIVE de-identification tool, and then get annotated. However, medical images are loaded already standardized directly with DICOM, so there are not as many standardization actions for medical images as there are needed for clinical data. For image annotation a series of workshops have taken place for the different types of cancer that INCISIVE deals with, and detailed guidelines on how to perform the annotations have been provided in previous Deliverable 4.1 INCISIVE AI-toolbox, data analytics and user services.

4.7.4.2. HL7 FHIR

HL7 FHIR[47] was chosen for clinical data exchange as it is a "RESTful" specification based on common industry-wide usage of the term REST. Considering that AI Engines need to query clinical data quickly and in a standardized way, the use of the RESTful methodology will make it much easier for AI Models and AI Services to query this data.

At INCISIVE it has been decided that the data upload process will be carried out using an Excel file that indicates the terms to be reported and that each Data Provider will upload to their Federated Node or in the Central Node anonymously indicating a patientID. This mechanism makes it possible to send data from different patients from the same data provider with the same file. Four file templates have been generated, one for each cancer type with different clinical data, for this reason it has been decided to create four different HL7 FHIR messages, one for each cancer type.

To properly understand the work that has been done, you must first know that HL7 FHIR is an interface specification that specifies the structure of the data to be exchanged

[45] DICOM Overview. (n.d.). from https://www.dicomstandard.org/current
[46] NIFTI Overview. (n.d.). from https://nifti.nimh.nih.gov/
[47] HL7 FHIR Overview. (n.d.). from https://www.hl7.org/fhir/index.html

between health applications/systems. It defines the following possible methods of data exchange:

- RESTful API
- Messaging
- Documents
- Services / SOA
- Database / Persistent Storage
- Subscriptions Framework

In INCISIVE it has been decided to use the RESTful API method[48] and the Messaging method[49], for more details you can consult the INCISIVE Interoperability framework.

Then the MIME-type "application/fhir+xml" has been chosen for the messages that will be designed, which indicates that the messages will be in XML format. Although JSON could also have been chosen, MIME was selected, because the initial documentation provided for implementation predominantly featured examples and templates in XML format. This influenced our decision to opt for XML to ensure consistency and ease of use. Further, once XML was selected, all subsequent test coverage was designed with this format in mind. Given the extensive efforts and resources already invested into establishing XML-based protocols, we decided against transitioning to JSON at a later stage to maintain continuity and reliability in our processes.

Following it has also been decided to use the FHIR R4 version ("4.0" mixed STU/Normative), as it was the last stable and normative version at the time of decision-making.

The rest of the implementation and configuration of the FHIR RESTful API and the FHIR Server is explained in the next Technical Interoperability section.

In this section we will focus more on the Messages and explain what has been done to define the message for each type of cancer.

Clinical data, transactions and semantic groupings have been analysed. It has been decided to process prospective data in the same way as retrospective data.

The data is uploaded anonymously to the template and the real reference is maintained through a patient-id. Therefore, the entire HL7 FHIR implementation guide of INCISIVE is based on anonymised data.

Regarding data groupings at a semantic level, the entire template is always loaded and all the data that is reported is saved. A guide has been defined with the cardinality of each

[48] RESTful API - FHIR v4.0.1. (n.d.). from https://hl7.org/fhir/R4/http.html
[49] Messaging - FHIR v4.0.1. (n.d.). from https://hl7.org/fhir/R4/messaging.html

term of the template to establish a common minimum of data to be reported whether they are prospective or retrospective (Added a Mandatory/Optional column on each message guide based on the FHIR specification and explained the use of Mandatory/Optional for different INCISIVE clinical data scenarios).

Term	ID resource	FHIR message location	Example Value	Required by FHIR specification	INCISVE Profile	Condition
Patient Number	patient01	Patient.identifier.value	004-000001	Optional		
Gender	patient01	Patient.gender	female	Optional		
Ethnicity	patient01	Patient.extension.valueCodeableConcept.coding.code	1	*Optional*	Extension	
		Patient.extension.valueCodeableConcept.coding.display	White	*Optional*		
Number of births	observation01	Observation.component.valueInteger	2	*Optional*		
Age at diagnosis	diagnosticReport01	DiagnosticReport.extension.valueAge.value	56	Mandatory	Extension	
Pre/post menopause	observation01	Observation.component.valueCodeableConcept.coding.code	0	*Optional*	CodeSystem	
		Observation.component.valueCodeableConcept.coding.display	Pre	*Optional*		
Medical History	condition01	Condition.code.text	N39.8	Optional		
Familial Cancer History Father	observation01	Observation.component.valueBoolean	true	*Optional*		

Figure 18 Breast cancer message guide example.

According to the different choices of Mandatory/Optional for terms, there will be different rules, which are also describe in the message guide.

2.3. Description mandatory/optional of term

There will be mandatory/optional for different scenarios, according to different scenarios corresponding to different rules. These scenarios are described separately below.

- If the attribute is mandatory, all attribute structure information for this column must be reported. It cannot be deleted.

Notes general info	observation01	**Observation.note.text**	Observation note	Mandatory

- If the attribute is mandatory with a comment reporting the condition of this variable, all attribute structure information for this column must be reported, unless the variable has not value for report, all resources must be deleted.

Clinical findings general info	condition03	**Condition.note.text**	Dysuria	Mandatory	If Clinical findings general info=null delete all condition03

- If an optional attribute needs to be deleted, all attribute structures except for the primary resource (e.g., observation, procedure, condition, etc.) must be deleted if that primary resource has another variable. For example, for the variable Gender, you can delete the attribute gender from the resource Patient, while other attributes such as identifier have values and do not have to be deleted.

Gender	patient01	**Patient.gender**	female	Optional

- If the attribute is optional, this happens in observation.component because Observation.component has two parts inside, Observation.component.code and Observation.component.value (this value can be valueCodeableConcept, ValueBoolean, ValueSrting, valueInteger, valueQuantity, etc.). If the variable has a value that needs to be informed, then code is mandatory, that code is mostly the SNOMED CT code or LOINC code for that variable. This means that if you need to delete variable, then you need to delete all Observation.component and cannot delete the whole Observation resource, unless there are no other elements in the same Observation, then you can delete the whole Observation resource.

Annotated Mammography baseline	observation02	**Observation.component.valueCodeableConcept.coding.code**	1	*Optional*	CodeSystem
		Observation.component.valueCodeableConcept.coding.display	Positive annotation	*Optional*	

- If the attribute is optional with a comment reporting the condition of this variable, you can delete all resource if the variable has not value for report.

Symptoms general info	condition02	**Condition.evidence.code.coding.code**	1	Optional	CodeSystem	If Symptoms general info=null delete all condition02

Figure 19 Description mandatory/optional of terms example.

On the other hand, before defining which resources are most suitable for each term, the groups of data from the template were analysed. It is very important to take these datasets into account when choosing the resources to know the limitations that will arise throughout the process in the specification of the elements for each term. It is also very important to consider the data loading processes and use cases to know if everything is loaded in a single message or if there will be more than one message for the same patient.

Considering that there are two clearly different data sets, we could have defined two Bundles, one for Prostate, and one for the other three types. But although the data sets between Breast, Colorectal and Lung are the same, the data to be exchanged within each dataset is different. For this reason, INCISIVE has decided to define a Bundle for each cancer and has decided to send it for each patient.

To do the data analysis and mapping with FHIR elements and decide which resources to use for each data set and the type of cancer, we have done the following actions:

- We have chosen the most suitable resource for each term according to the clinical meaning and the dataset to which it belongs.
- We have chosen the most appropriate element for each term according to the chosen resource. The "valueType" of each term has also been considered.
- All required local codes have been created to be used as the CodeSystem for each term.

- It has been analysed whether the elements and resources chosen to agree with the semantic system and the semantic tag of each term.
- With terms that could not be represented within the general structure of FHIR, their need has been analysed and the necessary FHIR extensions have been created for each resource.
- We have ensured that all FHIR bindings and restrictions are compliant with the proposed design.
- A conceptual map has been drawn up on how it has been decided to relate the different CodeSystems, with the ValueSets, Extensions and StructureDefinition.

The FHIR resources[50] that INCISIVE has decided to use are:

FHIR Resource	Scope and Usage	Count
Bundle	A container for a collection of resources.	5
Message Header	The header for a message exchange that is either requesting or responding to an action.	4
Organization	A group of people or organizations formed to achieve some form of collective action.	4
Patient	Demographics and other administrative information about an individual or animal receiving care or other health-related services.	4
Condition	A clinical condition, problem, diagnosis, or other event, situation, issue, or clinical concept that has risen to a level of concern.	10
Observation	Measurements and simple assertions made about a patient, device or other subject.	46
Diagnostic Report	The findings and interpretation of diagnostic tests performed on patients, groups of patients, devices, and locations, and/or specimens derived from these.	4
Procedure	An action that is or was performed on or for a patient. This can be a physical intervention like an operation, or less invasive like long term services, counselling, or hypnotherapy.	31
Medication	The identification and definition of a medication for the purposes of prescribing, dispensing, and administering a medication as well as for making statements about medication use.	4
Medication Administration	Describes the event of a patient consuming or otherwise being administered a medication.	5

[50] HL7 FHIR Resourcelist - FHIR v4.0.1. (n.d.). from https://hl7.org/fhir/R4/resourcelist.html

CodeSystems	Used to declare the existence of and describe a code system or code system supplement and its key properties, and optionally define a part or all its content.	67
ValueSets	Specifies a set of codes drawn from one or more code systems, intended for use in a particular context. Value sets link between definitions and their use in .	13
Extensions	Every resource or datatype element may include one or more "extension" child elements. The extension is either simple or complex. Simple extensions have only a value and no nested extensions. Complex extensions contain one or more nested extensions and no value. An extension cannot have both a value and nested extensions. Every extension in a resource refers directly to its definition, which is made available as a StructureDefinition. A resource can be profiled to specify where extensions are required or expected.	3
ElementDefinition	The definition of an element in a resource or an extension. ElementDefinition is used in StructureDefinition.	9
StructureDefinition	Describes a structure - a set of data element definitions, and their associated rules of usage. These structure definitions are used to describe both the content defined in the FHIR specification itself - Resources, data types, the underlying infrastructural types, and are used to describe how these structures are used in implementations. This allows the definitions of the structures to be shared and published through repositories of structure definitions, compared with each other, and used as the basis for code, report and UI generation.	9
CapabilityStatement	Documents a set of capabilities (behaviours) of a FHIR Server for a particular version of FHIR that may be used as a statement of actual server functionality or a statement of required or desired server implementation.	1
ImplementationGuide	A set of rules of how a particular interoperability or standards problem is solved - typically using FHIR resources. This resource is used to gather all the parts of an implementation guide into a logical whole and to publish a computable definition of all the parts.	5

Table 11 FHIR resource count.

*Blue background: resources used to represent the terms of the template.

The diagram explaining the terminological design of HL7 FHIR that has been decided to be used in INCISIVE can be consulted in the INCISIVE Interoperability framework.

Based on the terminology design diagram, INCISIVE has created corresponding FHIR resources to fit the specific usage context of the project. FHIR resources are designed with the 80/20 rule in mind - focus on the 20% of requirements that satisfy 80% of the interoperability needs. To this end, resources are designed to meet the general or

common data requirements of many use cases. On this basis, FHIR allows for the adoption of generic resources that are adapted to specific use case requirements.

As a result, INCISIVE was able to personalize and restrict the required FHIR resources according to different use cases.

FHIR allows designers to specify:

- Rules about which resource elements are or are not used (through StructureDefinition), and what additional elements (in form of extensions) are added that are not part of the base specification.
- Rules about which of FHIR's RESTful API, messaging and document features are used, and how (using Capability Statement).
- Rules about which terminologies are used in each element (using ValueSet & CodeSystem resources).
- Descriptions of how the Resource elements and API features map to local requirements and/or implementations (using ImplementationGuide).

These are the CodeSystem created for INCISIVE:

Type of cancer	CodeSystem.name	
All cancer	• INCISIVE Data Providers • Current state of patient • Ethnicity of patient	• Label Timepoints • Label Treatment • Label Exam • Label Biopsy
Breast cancer	• Type of Surgery Breast • Type of Biopsy Breast • Symptoms General Info Breast • SymptomsTimepoints Breast • Breast Cancer Type • Molecular Subtype Breast • Baseline location Breast • Distant metastasis location Breast • Annotated DICOM files Breast • Response to treatment Breast • Lymph-node surgery Breast • Pre/post menopause Breast • Breast laterality • Mass Shape Breast	• Mass Margin Breast • Mass orientation Breast • Mass echo pattern Breast • Posterior features Breast • Calcifications baseline Breast • Associated Features Breast • Breast composition • Parenchymal enhancement level Breast • Parenchymal enhancement symmetry Breast • Presence of foci Breast • Breast BIRADS classification • Calcifications timepoints Breast
Colorectal cancer	• Syndromes Colorectal • Location Colorectal • Location for rectal cancer Colorectal	• Distant metastasis location Colorectal • Annotated DICOM files Colorectal • Response to treatment Colorectal

	• Location of RT Colorectal	
Lung cancer	• Lung Cancer Type • Lung nodules location • Distant metastasis location baseline Lung • Distant metastasis location timepoints Lung • Annotated DICOM files Lung	• Response to treatment Lung • Smoker Lung • Laterality Lung • Lobe Lung • Solidity Lung • Borders Lung
Prostate	• Type of therapy Prostate • Type of Biopsy Prostate • Symptoms General Info Prostate • Colon Cancer Type • Localization of PIRADS Prostate • Treatment Prostate • Annotated DICOM files Prostate • Response to treatment Prostate	• MRI type baseline Prostate • Side of tumour Prostate • Maximum PIRADS Prostate • pT histology-mutations Prostate • pN histology-mutations Prostate • cT Baseline of Prostate Cancer • cN Baseline of Prostate Cancer • cM Baseline of Prostate Cancer

Table 12 INCISIVE CodeSystem.

A total of 67 CodeSystem have been created. There were 7 CodeSystem that can be shared for 4 cancers, 28 for breast cancer, 11 for lung cancer, 13 for prostate cancer and 8 for colorectal cancer.

A total of 13 ValueSet has been created. All of them contain more than one CodeSystem. Five elements of the VauleSet can be shared across four cancers, three can be shared with several of the cancers but not all cancers, and five are used only for specific cancers.

In summary, INCISIVE has been profiled a total of 11 StructureDefinition, which are profiled for the Organization, Patient, Procedure, Condition, DiagonisticReport, and Observation resources. Of these, the Procedure and Observation resources each define multiple StructureDefinition for different cancers. Several other resources can be shared across the four cancers.

INCISVE decided to create a primary ImplementationGuide that would contain four different ImplementationGuide for each of the four cancers. This will make it easier for future implementers who want an implementation guide for a particular cancer to simply reuse the ImplementationGuide for that cancer, avoiding redundant data.

Each of the four ImplementationGuide consists of FHIR resources and human-readable documents. Resources contain the following:

- Terminology: CodeSystem and ValueSet
- Content rules: Profiles and Extensions
- API Details: CapabilityStatement

- Supporting Examples

The human-readable documents contain:

- Describe security arrangements in place for the relevant data exchanges, along with consent and access control requirements.
- Detail additional rules about the user of the narrative in the resources.

All resources created are published on Simplifier[51] and GitHub. GitHub[52] has been used to store all XML files, it has file version control, that can see every change on the file. The Terminology resources creation and publication has been done with Forge and Simplifier. Forge has been used to design the terminologies and Simplifier to publish them.

In addition, all resources are also validated using the FHIR validator[53] to ensure that the resources are structured correctly and comply with the FHIR specification.

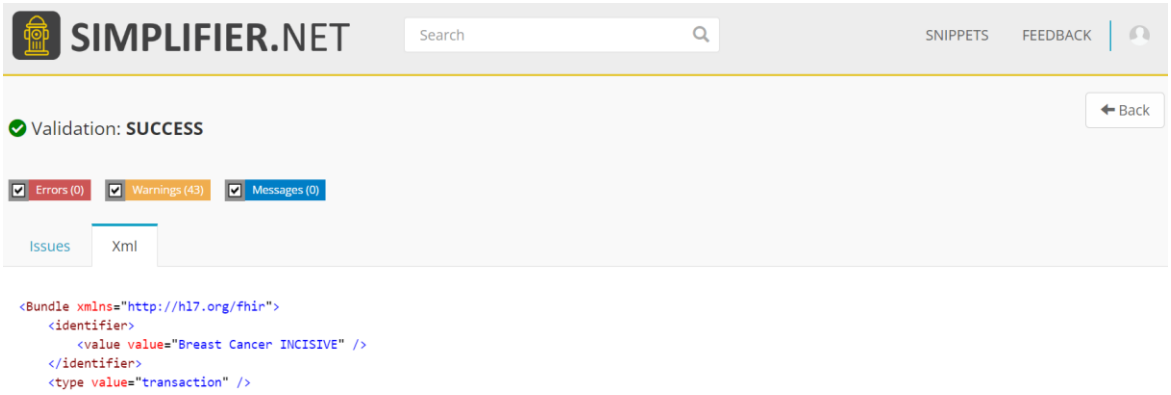


Figure 20 Breast bundle validation.

The Simplifier's FHIR validator is used in INCISIVE, but there are other validators on the market that could work just as well.

Forge[54] is an application for authoring, creating, editing, validating FHIR profiles, extensions, and implementation guides.

[51] The FHIR collaboration platform - SIMPLIFIER.NET Overview. (n.d.). from https://simplifier.net/
[52] GitHub Overview (n.d.). from https://github.com/
[53] Validator FHIR. (n.d.). from https://simplifier.net/validate?fhirVersion=R4&scope=hl7.fhir.r4.core@4.0.1
[54] Forge Overview. from https://fire.ly/products/forge/

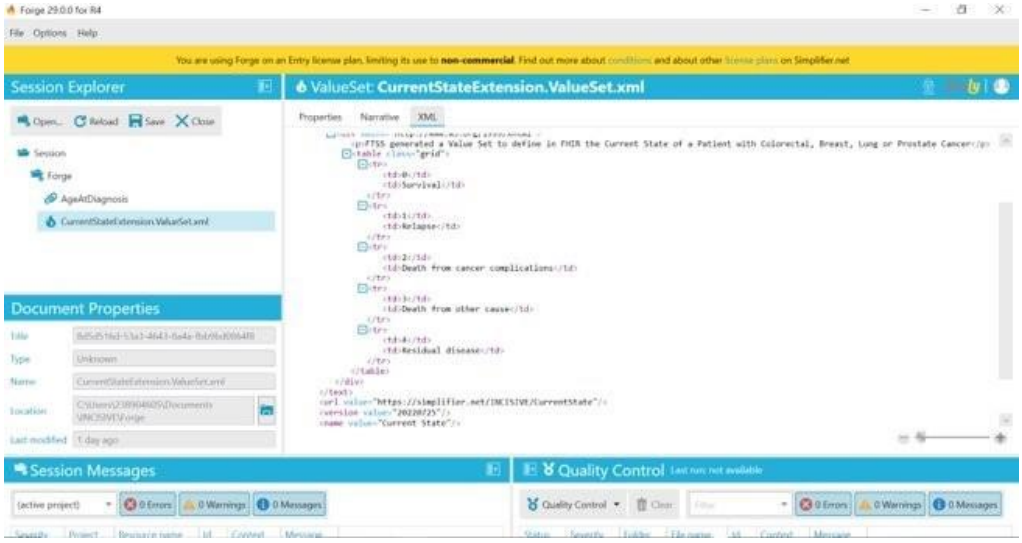


Figure 21 Capture of Forge.

Forge is integrated with Simplifier; it is possible to fetch and publish resources to the Simplifier projects.

Simplifier is a FHIR collaboration and publishing platform, where it is possible to create, upload, download, find a view FHIR resources.

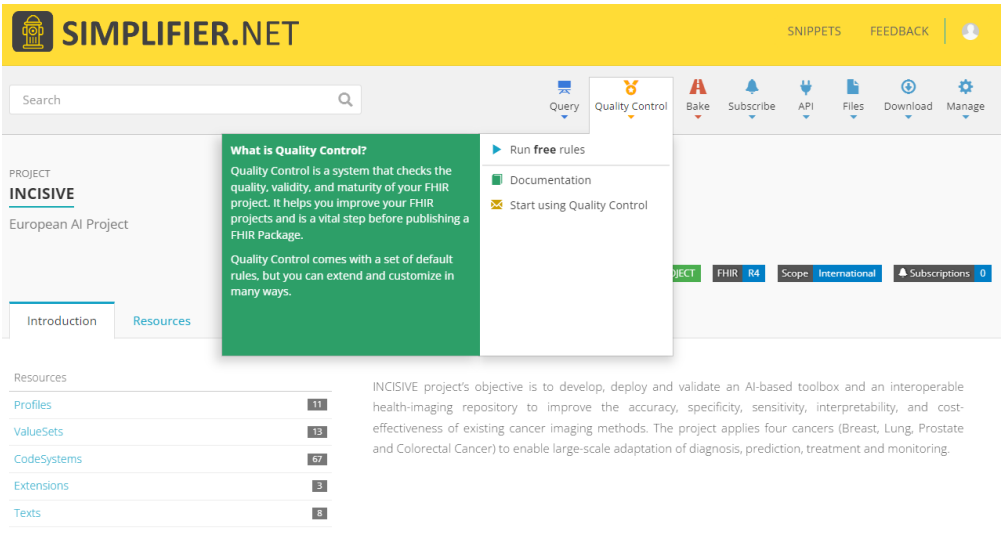


Figure 22 Capture of INCISIVE project in Simplifier.

The publication and consultation of the implementation guide is explained in section 5.3 of this document.

The Bundle of each cancer defined for INCISIVE can be consulted in Annex 2 of this document. However, the Bundle Terminology/Server config can be only consulted in the INCISIVE Interoperability framework.

Bundle	Number of Entries	Lines per message
Breast	30	3242
Lung	31	2795
Prostate	19	2229
Colorectal	25	2451

Table 13 Bundle count.

4.7.4.3. Clinical Report and uploading data for model inference

As described in Section 4.3 for model inference use case 1 and 2, the INCISIVE platform provides a variety of AI services to healthcare professionals that generate the results of AI models. Each cancer type has their AI service, and each AI service corresponds pipeline that use different AI models.

As an example, the figure below shows the AI service for breast cancer: identifying the AI models with box, the clinical data with database and medical images with the image modalities type allows for this AI service. The box in the middle of each AI model represents an algorithm, and each algorithm generates corresponding results which are shown in the top of the box. All these results of AI models are collected on the final clinical report.

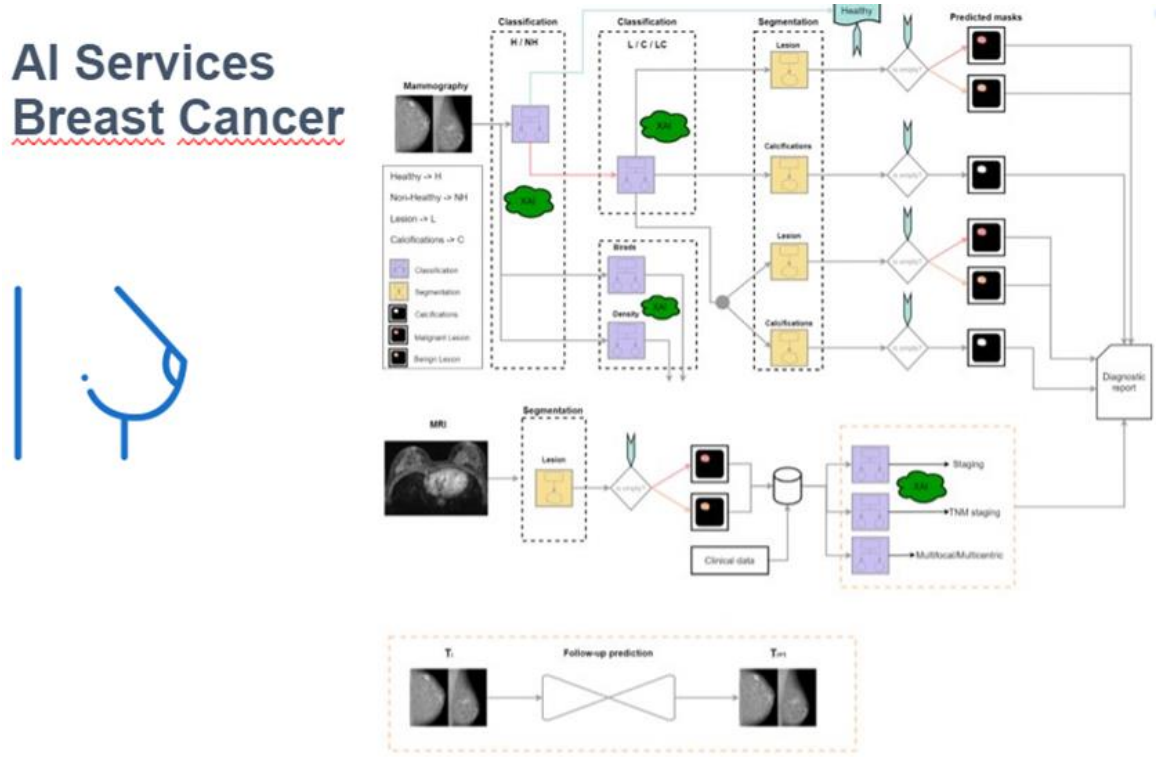


Figure 23 Breast Cancer AI Service.

Following the medical terminology, the required parameters that the clinical report should encompass (under an automatic or at least with a minimal assistance way) are parameters that describe the visual content of the corresponding medical image. For example, the name of the tumour, the location of the tumour, the organs affecting, the size of the tumour, its metastatic behaviours, etc. Each one of the above parameters is considered an individual field of information, which is correlated with a corresponding descriptive medical sentence. On the top of that the clinical report should also include a metadata section that presents information about the study, the patient and the service that is running. That kind of metadata should be institute name, patient ID, clinical report date, etc.

The sources that have been used to generate the definitions and the terminology standards for each cancer type are national and world-renowned medical research organizations in the fields of breast, lung, colorectal and prostate cancer, such as the National Cancer Institute of the US, the Cancer Research of UK, the NHS of UK, and the American Cancer Society, following medical standards as AJCC[55]. On top of that, the

[55] AJCC staging system - NCI Dictionary of Cancer Terms - NCI. (n.d.). from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/ajcc-staging-system

generated reports are getting reviewed by the healthcare professionals of INCISIVE, to ensure their validity from the medical point of view and to be compliant with the requirements of the INCISIVE project.

HL7 CDA[56] will ensures interoperability and consistency in the exchange of clinical information and the PDF report (a user-friendly layout with clear headings, sections, and visual elements) will facilitate easy comprehension by healthcare professionals.

The following steps have been done to standardize the medical report template for each cancer:

1. Organize and tag the results generated by the AI model.
   The possible results generated by the AI model have been collected in an Excel sheet for each AI service of each cancer to elaborate the PDF template and HL7 CDA template. Refer to the following screenshot, the Excel contains 6 columns:

Column	Description
Report Visualization Order	Indicate the order in which each term is displayed in the clinical report template.
Term Name of each box show in the above image	These terms include information such as clinical report type, patient and healthcare professional, and various algorithms in the pipeline. By combining the values of these terms, a complete clinical report can be generated.
Term Code	Each term has a corresponding abbreviation code for ease of reference by the technical partner.
Possible Values	The possible values are the possible algorithm results of each AI model.
Possible Values Code	Are code for identifying possible values.
Sentences	The sentence is the human-readable translation of the possible value result of each AI model and are the ones that will be include PDF template and HL7 CDA template.

Table 14 Legend Terms Clinical Report.

In the clinical report template, there are seven terms that are common across all four cancers: Tumour name/Organs affecting, patient ID, Healthcare professional ID, Clinical report ID, Date of report generated, Data provider and Imaging modality type. The other terms vary depending on the type of cancer, and these terms correspond to different algorithms.

[56] HL7 CDA Overview. (n.d.). from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

Report Visualization Order	Variable Name	Variable Code	Possible Values	Possible Values Code	Sentences
1	Tumor Name/Organs Affecting	[Organ]	Breast Cancer	[Organ:Breast]	Report for breast cancer.
			Colorectal Cancer	[Organ:Colorectal]	Report for colorectal cancer.
			Lung Cancer	[Organ:Lung]	Report for lung cancer.
			Prostate Cancer	[Organ:Prostate]	Report for prostate cancer.
	IdPatient	[Patient ID]			005-00000
	IdHCP	[HCP ID]			UNS0000
	ReportId	[Report ID]			BC000
	Generated Report Date	[Date Created - YYYYMMDD]			YYYYMMDD
3	Data Provider	[Institute NAME]	Name of HCP Institute	N/A	i.e. UNS.
2	Imaging Modality Type	[Modalities]	Modality MRI	[Modalities: MRI]	Examination regards MRI uploaded modalities.
			Modality Mammography	[Modalities: M]	Examination regards Mammography uploaded modalities.
4	Healthy/Not Healthy	[H/NH]	H	[H/NH:H]	TextH
			NH	[H/NH:NH]	The breasts appear to contain oncological findings for breast cancer with an abnormal mass or nodule.
5	BIRADS	[BIR]	BIRADS 1	[BIR:1]	Text1
			BIRADS 2	[BIR:2]	Text2
			BIRADS 3	[BIR:3]	Text3
			BIRADS 4	[BIR:4]	Suspicious abnormality/malignacy has been found. These findings do not definitely look like cancer but could be cancer. A biopsy should be considered.
			BIRADS 5	[BIR:5]	Text5
			BIRADS 6	[BIR:6]	Text6
6	Density	[DENSITY]	DENSITY 1	[DENSITY:1]	Text1
			DENSITY 2	[DENSITY:2]	Text2
			DENSITY 3	[DENSITY:3]	More of the breast is made of dense glandular and fibrous tissue (described as heterogeneously dense). This can make it hard to see small masses in or around the dense tissue, which also appear as white areas. Heterogeneously dense indicates that there are some areas of non-dense tissue, but that the majority of the breast tissue is dense.
			DENSITY 4	[DENSITY:4]	Text4
8	Staging	[Stage]	STAGE 1	[Stage:1]	Text1
			STAGE 2	[Stage:2]	The tumour size should be between 2-5 cm, inside the breast with or without spread to the axillary lymph nodes. Or the tumour is 2-5 cm and the disease has spread to 1 to 3 axillary lymph nodes. Or the tumour is larger than 5 cm but has not spread to the axillary lymph nodes. This is an early stage breast cancer.
			STAGE 3	[Stage:3]	Text3
			STAGE 4	[Stage:4]	Text4
9	Staging TNM - Size(T)	[T]	T1	[T:T1]	TextT1
			T2	[T:T2]	The recommendation is for T2, meaning that the tumour is more than 2 centimetres but no more than 5 centimetres across.
			T3	[T:T3]	TextT2
			T4	[T:T4]	TextT3
10	Staging TNM - Invasiveness(N)	[N]	N0	[N:1]	Text1
			N1	[N:2]	The cancer cells have spread to one or more lymph nodes in the lower and middle part of the armpit. The lymph nodes move a little when they are felt and are not stuck to surrounding tissue.
			N2	[N:3]	Text3
			N3	[N:4]	Text4
11	Staging TNM - Metastasis(M)	[M]	M0	[M:0]	There are no sign of metastasis to other organs, meaning that the cancer has not spread to another part of the body.
			M1	[M:1]	Text1
7	Location	[LOC]	UIQ	[LOC: UIQ]	TextUIQ
			UOQ	[LOC: UOQ]	The tumour appears to be concentrated to at the upper outer quadrant of the breast.
			LIQ	[LOC: LIQ]	TextLIQ
			LOQ	[LOC: LOQ]	TextLOQ

Breast | Colorectal | Lung | Prostate | Common Variable Cancer | Example Breast text MRI (PDF)

Figure 24 Clinical Report Excel Template.

2. Generate four clinical report templates in PDF and HL7 CDA format.
   After the operation of the clinical report inside the INCISIVE platform, the service outputs a report in XML format also including metadata, and additionally a report in PDF format, to be delivered to the user through the front-end.
   - The PDF format template
     The screenshot is an example of the PDF template for the clinical report. The first page contains common terms, and the second page shows the results generated by the AI models. In each bracket are the contents to be filled, each represented by a term code, which can be found in the Excel sheet above. And the filled content is filled according to the results of AI models.

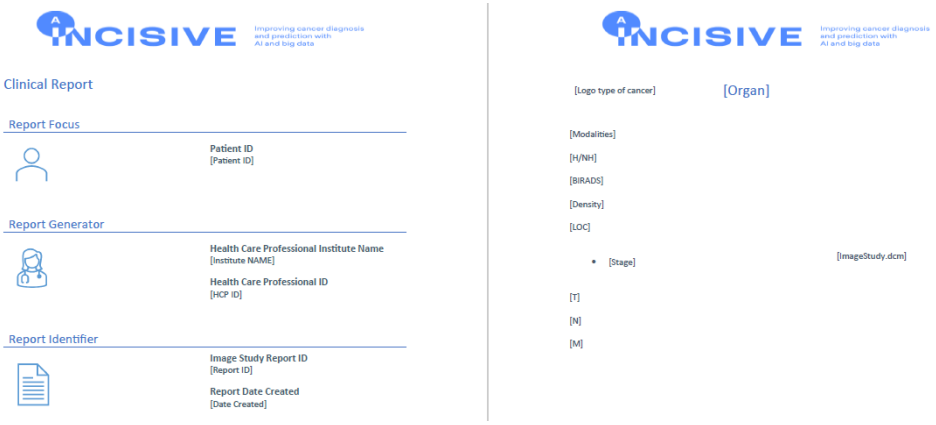


Figure 25 Clinical report PDF template.

As an example, this will be a clinical report for breast cancer AI Service.

Figure 26 Breast cancer clinical report example.

- The HL7 CDA format template
  HL7 CDA is a document markup standard that specifies the structure and semantics of "clinical documents". CDA is an XML encoded metadata fields (such as provider name, document type, document identifier, and so on) and a body which can be any commonly used MIME type such as pdf or .doc or a structured data element. The body presents the information it contains in a human-readable format, but it also carries the information in data structures that are machine processable. It has a header and a body:
  The purpose of the CDA header is to set the context for the report. The INCISIVE HL7 CDA header in template contains the common terms such as Tumour name, patient and healthcare professional id, clinical document id and

generated date, DataProvider name, etc. in the clinical report, and each term is represented in the appropriate ClinicalDocument's attribute.

```
**********************************************************
  CDA Header
**********************************************************
-->
    <typeId root="2.16.840.1.113883.1.3" extension="POCD_HD000040"/>
    <templateId root="2.16.840.1.113883.3.27.1776"/>
    <!--Clinical Report ID-->
    <id extension="[Report ID]"/>
    <code code="371524004" codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT" displayName="Clinical report (record artifact)"/>
    <!--Tumor Name/Organs Affecting-->
    <title>[Organ]</title>
    <!--Report Date Created-->
    <effectiveTime value="[Date Created - YYYYMMDD]"/>
    <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25"/>
    <languageCode code="en-US"/>
    <setId extension="BB35" root="2.16.840.1.113883.19.7"/>
    <versionNumber value="2"/>
    <recordTarget>
        <patientRole>
            <!--Patient ID-->
            <id extension="[Patient ID]"/>
        </patientRole>
    </recordTarget>
    <author>
        <!--Report Date Created-->
        <time value="[Date Created - YYYYMMDD]"/>
        <assignedAuthor>
            <!--Health Care Professional ID-->
            <id extension="[HCP ID]"/>
            <representedOrganization>
                <!--Health Care Professional Institute Name-->
                <id extension="[Institute NAME]"/>
            </representedOrganization>
        </assignedAuthor>
    </author>
    <custodian>
        <assignedCustodian>
            <representedCustodianOrganization>
                <id extension="XXX"/>
                <name>Plataforma INCISIVE</name>
            </representedCustodianOrganization>
        </assignedCustodian>
    </custodian>
```

Figure 27 CDA body template.

The CDA body contains the results generated by the AI models and can be an unstructured block, represented by the NomXMLBody class, or it can be composed of a structured data element, represented by the StructuredBody class. INCISIVE have two types of body designed, but only allows to download HL7 CDA with structure data element body to be more standardized possible. The features of both are described to follow:

- **Unstructured block PDF B64**
  The NonXMLBody class is a simple way to wrap an existing non-XML document with a CDA header, such as wrapping PDF B64. The step is to translate the PDF clinical report template into PDF format using a converter, and then add it into the nonXMLBody element; there is no encryption. But in the future the actual use of the need to take the encryption method. This way represents a very low standard for adoption of the standard.
- **Structure data element**
  The StructuredBody contains one or more Section components. Each term of the box shown above is an object of a CDA element, represented using the entryRelationship element, rather than being converted to an associated PDF template. The method is intended for making it incrementally interoperable so that it can be interpreted and manipulated by the machine.

All the details about the definition process of the PDF templates and the HL7 CDA templates are explained in the Interoperability Framework.

Regarding the upload of the data to make use of the services that the INCISIVE toolbox provides, this input is provided through the UI. To be as simple as possible to the healthcare professional, the input is defined as a zip file containing the images to be analysed in DICOM format along with the required tabular/textual data.

Notice that the current solution requires to put the tabular information in a defined format inside the zip file, however in the future this part should be provided as part as the UI in the form of textboxes that check the information input and format it in a standard way inside the platform.

4.7.5. INCISIVE Technical Interoperability

FHIR Server was chosen for clinical data considering that AI Engines need to query this data quickly and in a standardized way, and PACS[57] was chosen for medical image and annotations to process the DICOM files; the use of the RESTful methodology will make it much easier for AI Models and AI Services to query this data.

The architecture plan of INCISIVE has a lot to do with the decisions that have been made when choosing the technology to use. A hybrid, federated and centralized system has been proposed.

This means that data for training models can be loaded from a federated node as a data provider, or from a centralized node that acts as a federated node (Is only available to any entity that wants to provide data and do not have the infrastructure to create a federated node).

Data from the inference models is loaded into the central node directly from the INCISIVE UI and is currently non-persistent, the data is consumed, and results are generated at runtime for each healthcare professional.

The technology that has been used to support this federated and scalable system has been through microservices that can grow horizontally within the system, as well as adding different services that the platform wants to offer.

Some of the services used by the platform for its operation are:

[57] PACS Overview. Arora D, Mehta Y. Use of picture archiving and communication system for imaging of radiological films in cardiac surgical intensive care unit. J Anaesthesiol Clin Pharmacol. 2014 Jul;30(3):447-8. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4152706/

- **Orthanc PACS**: This is a comprehensive server dedicated to the handling of DICOM files, facilitating the storage, retrieval, and distribution of medical imaging information.
- **IBM FHIR Server**: This server aligns with the HL7 FHIR specifications, enabling the consistent representation and exchange of healthcare-related data. Its primary role is to facilitate interoperability between data sources, systems, and applications within the healthcare domain.
- **ETL Tool**: It is an internal tool that allows you to Extract then Transform Excel templates with clinical data into HL7 FHIR messages and finally Load this transform data in the FHIR server. It is also responsible for keeping the medical images and annotations in the PACS.
- **Quality Check Tool**: It is a tool developed to ensure that DICOM files and Excel templates have the correct format and data types before uploading them to the federated node or centralized node. It is a tool that runs locally on the main terminal of the data provider.
- **De-Identification Tool**: It is a tool developed to facilitate anonymisation of the medical images before uploading them to the federated node or centralized node. It is a tool that runs locally on the main terminal of the data provider.
- **Transaction Tracker Tool**: It is an internal tool developed to save the actions carried out on the INCISIVE platform, it performs the audit using blockchain. It is completely transparent to the user, is an internal tool of the architecture available to the auditor.

4.7.5.1. PACS

For INCISIVE, we selected Orthanc PACS to store medical images. Orthanc PACS[58] is a free, open source, lightweight DICOM server designed specifically for handling medical imaging data. Its architecture is light and self-contained, negating the need for complex database administration or the installation of third-party dependencies. Furthermore, Orthanc PACS provides a fast and lightweight RESTful API, which facilitates scalability and seamless integration into our system.

The setup and configuration of Orthanc PACS within our system was accomplished using a Docker file, which allowed for a streamlined and reproducible setup process.

This Docker file creates a container from the osimis/orthanc image and copies the orthanc.json configuration file into the appropriate directory within the container. This configuration file includes several key elements:

- **Name** specifies the name of our PACS system.
- **AuthenticationEnabled** and **RegisteredUsers** are for user authentication.

[58] Orthanc PACS. (n.d.). from https://www.orthanc-server.com/

- **PostgreSQL** contains database connection details, enabling indexing and storage. This section includes the database name, host, port, and access credentials (username and password).
- **UserMetadata** is our provision for custom metadata storage. While this feature is currently used experimentally, we foresee its active utilization in future improvements to our system.

Regarding the username and password under the PostgreSQL section, it's important to note that these values are placeholders and are overridden at runtime with the actual credentials. This practice enhances the security of our system by preventing sensitive access information from being hard coded into the configuration file.

The table below contains some examples of how some searches are performed technically, using API REST parameters of the Orthanc PACS:

QUERYS	
Upload a new DICOM file (Image-01.dcm)	pacsuser:incisive -X POST http://incisive-dp-pacs:XXXX/instances --data-binary @Image-01.dcm
List all patients IDs	pacsuser:incisive http://incisive-dp-pacs:XXXX/patients
Get information about a patient (by the patient ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/patients/{PatientID}
Delete a patient (by the patient id)	pacsuser:incisive -X DELETE http://incisive-dp-pacs:XXXX/patients/{PatientID}
List all studies IDs	pacsuser:incisive http://incisive-dp-pacs:XXXX/studies
List all DICOM instances IDs	pacsuser:incisive http://incisive-dp-pacs:XXXX /instances
Get all studies of a specific patient (by patient ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/patients/{PatientID}/studies
Get detailed information about a specific study (by study ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/studies/{StudyID}
Get all series of a specific study (by study ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/studies/{StudyID}/series
Get detailed information about a series (by series ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/series/{SeriesID}
Get all instances of a specific series (by series ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/series/{SeriesID}/instances
Get detailed information about a specific DICOm instance (by instance ID)	pacsuser:incisive http://incisive-dp-pacs:XXXX/instances/{InstanceID}
Find series with modality MG with detailed information	pacsuser:incisive -X POST http://incisive-dp-pacs:XXXX/tools/find -d '{"Expand": true, "Level": "Series", "Query": {"Modality": "MG"}}'

Find all instances with detailed information of a Patient (PatientID)	pacsuser:incisive -X POST http://incisive-dp-pacs:XXXX/tools/find -d '{"Expand": true, "Level": "Instances", "Query": {"PatientID": "PatientID"}}'

Figure 28 INCISIVE Orthanc PACS queries.

Moreover, with our Orthanc PACS setup, we've unlocked the full range of functionalities offered by the Orthanc REST API. This comprehensive API facilitates versatile manipulation of the stored medical imaging data, including retrieval, updating, and deletion operations. For further details on the extensive capabilities of this REST API, you can refer to the official Orthanc REST API documentation[59].

4.7.5.2. FHIR Server

For INCISIVE, the IBM FHIR Server[60] has been chosen to store clinical data, as it is an open-source Java solution that supports the processing, validation, and storage of healthcare data according to the HL7 FHIR specification. This decision was also supported by the fact, that the ETL Tool is also a Java solution as well as the IBM is a global company that will provide maintenance support.

To configure the IBM FHIR Server a CapabilityStatement and a Bundle resource has been created with this resources and related Terminology resources to customize the FHIR server of INCISIVE with the defined specification.

A guide for each HL7 FHIR message has also been elaborated for the developers, with the cardinality of each term and with helpful comments to implement the ETL Tool to transform the data from each Excel template into the defined HL7 FHIR messages.

The searches to be carried out for the training models have also been defined. From the search section of the INCISIVE platform, searches can be made on all federated nodes, that is, on each FHIR and PACS server of all available nodes.

The table below contains some examples of how these searches are performed technically, using FHIR parameters and the FHIR RESTful API. The full specification of the queries can be consulted in the INCISIVE Interoperability framework:

QUERYS	URL_QUERY_EXAMPLE_XTypeCancer
Patients with [X type] cancer	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer]

[59] Orthanc REST API. (n.d.). from https://api.orthanc-server.com/
[60] IBM FHIR Server. (n.d.). from https://www.ibm.com/downloads/cas/EWNW6QE3?mhsrc=ibmsearch_a&mhq=FHIR%20Server

Patients with colorectal cancer	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000
Patients with [X type] cancer who are [female/male]	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer]&gender=[female/male]
Patients with colorectal cancer who are female	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&gender=female
Patients with [X type] cancer and age at diagnosis between [minimum value-maximum value] years	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:DiagnosticReport:patient:age-at-diagnosis=ge[minimum value]&_has:DiagnosticReport:patient:age-at-diagnosis=le[maximum value]
Patients with colorectal cancer and age at diagnosis between 30-50 years	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:DiagnosticReport:patient:age-at-diagnosis=ge30&_has:DiagnosticReport:patient:age-at-diagnosis=le50
Patients with [X type] cancer and have familial cancer history	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:Observation:patient:component-code-value-boolean=275937001:408732007=66839005$true,275937001:408732007=72705000$true
Patients with colorectal cancer and have familial cancer history	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Observation:patient:component-code-value-boolean=275937001:408732007=66839005$true,275937001:408732007=72705000$true
Patients with [X type] cancer and those who had done CT and histopathology image	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:Procedure:patient:code=77477000&_has:Procedure:patient:code=394597005
Patients with colorectal cancer and those who had done CT and histopathology image	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Procedure:patient:code=77477000&_has:Procedure:patient:code=394597005
Patients with [X type] cancer and those who had done CRT and CT	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:Observation:patient:component-code=77477000&_has:Observation:patient:component-code=252416005
Patients with colorectal cancer and those who had done CRT and CT	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Observation:patient:component-code=77477000&_has:Observation:patient:component-code=252416005
Patients with [X type] cancer and specific date timepoints	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:Observation:patient:category-text-value-integer=Label+timepoints$ge11
Patients with colorectal cancer and specific date timepoints	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Observation:patient:category-text-value-integer=Label+timepoints$ge11

Patients with [X type] cancer and those who had done CT and MRI	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of X type cancer] &_has:Observation:patient:component-code=77477000&_has:Observation:patient:component-code=113091000
Patients with breast cancer and those who had done CT and MRI	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Observation:patient:component-code=77477000&_has:Observation:patient:component-code=113091000
Patients with [colorectal/lung] cancer and specific cancer stage (stage 0, I, IIA, IIB, IIC, IIIA, IIIB, IIIC, IVA, IVB)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of colorectal/lung cancer] &_has:Observation:patient:component-code-value-string=385356007$[stage code]
Patients with colorectal cancer and specific cancer stage (stage IIA)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=781382000&_has:Observation:patient:component-code-value-string=385356007$IIA
Patients with lung cancer and specific stage (stage IA3)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=93880001&_has:Observation:patient:component-code-value-string=385356007$IA3
Patients with [breast] cancer and specific BIRADS classification (BIRADS 0, 1, 2, 3, 4, 5, 6)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of breast cancer] &_has:Observation:patient:component-code-value-concept=254292007$BI-RADS+[BIRADS number]
Patients with breast cancer and specific BIRADS classification (BIRADS 3)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=254837009&_has:Observation:patient:component-code-value-concept=254292007$BI-RADS+3
Patients with [prostate] cancer and specific Max PIRADS classification (PIRADS 1, 2, 3, 4, 5)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=[SNOMED CT code of prostate cancer] &_has:Observation:patient:component-code-value-concept=254292007$[PIRADS number]
Patients with prostate cancer and specific Max PIRADS classification (PIRADS 1)	serverBaseR4/Patient?_count=10000&_has:Observation:patient:code=399068003&_has:Observation:patient:component-code-value-concept=254292007$1

Table 15 INCISIVE FHIR Server queries.

At the same time, the path where the image of a patient id is stored, has been added in one FHIR resource of the same patient id HL7 FHIR message. In this way, with a query to the FHIR server, you get the path where the image related to the considered patient is located. With this path, the PACS server can be queried to access the specific image.

4.7.5.3. ETL Tool

The Extraction-Transformation-Load (ETL) component functionality allows a Data Provider (DP), in a Federated or a Central Node, to import its local clinical data into the INCISIVE platform. The INCISIVE ETL procedures have been designed using open standards based on LOINC, SNOMED CT, FHIR and PACS. LOINC and SNOMED CT are related to the Common Data Model and link to the Data semantic whereas FHIR is linked to the interoperability in

terms of data structure. Using well-known open standards allows current INCISIVE DPs to be incorporated into INCISIVE following a common, uniform way, effectively saving time and effort. An additional advantage of using open standards is that future DPs – currently not associated with INCISIVE – can be onboarded into INCISIVE with relatively small effort and low complexity.

In the INCISIVE platform, the Extraction-Transformation-Load (ETL) tool plays a pivotal role in integrating clinical data. The data collection process starts with a Microsoft Excel template, each line of which represents an individual subject. This Excel file is placed in a specific directory monitored continuously by the ETL tool.

The Federated Nodes of each Data Provider and the storage volumes used in the Central Node adopt a similar directory structure, patterned as /data/incisive-dp-{ORG}, where {ORG} represents the Data Provider. This structure facilitates efficient data ingestion by the ETL tool and the Federated Storage mechanisms.

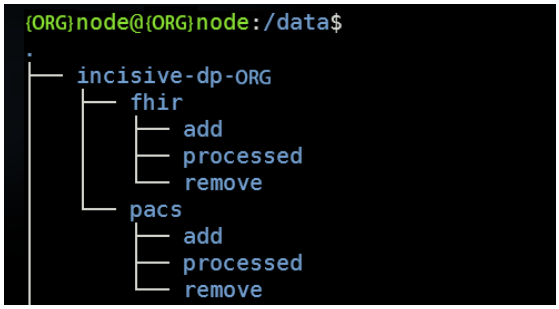


Figure 29 Federated Storage directory structure.

Within this directory structure, specific subdirectories are designated for different data ingestion tasks. For instance, 'fhir/add' and 'pacs/add' are the folders where Excel files and patient folders containing DICOM and NIFTI images, respectively, are placed for ingestion into the FHIR and PACS servers. Following the completion of data ingestion, the ETL Tool categorizes the processed files into 'success' and 'error' subdirectories within 'fhir/processed' and 'pacs/processed', indicating the outcome of the ingestion process.

- For Medical Images

  Before ingesting data into the PACS server, the data must first be standardized by the data-quality check tool, which generates output in a specific structure, typically in the format: /{patient}/{patient-study-timeline}/Series-{number}/Untitled.nii.gz and /{patient}/{patient-study-timeline}/Series-{number}/image-1.dcm.

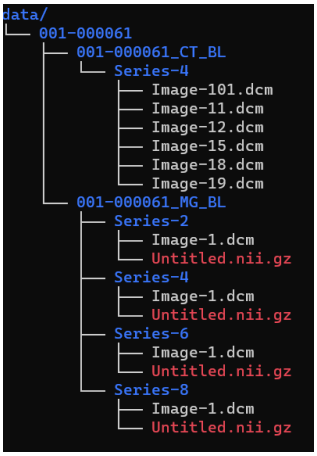


Figure 30 PACS server structure.

Data ingestion into the PACS server involves moving the standardized patient directories into the 'pacs/add' subdirectory. For every DICOM image, the ETL tool sends a POST request to the specific server endpoint at http://incisive-dp-pacs:XXXX/instances. This process can be time-consuming given the size and quantity of these files and folders, but the ETL tool then automatically begins processing and ingesting the data.

- For Clinical data

  The ingestion process for the FHIR server mirrors that of the PACS server, with a slight difference. Instead of moving patient directories, Excel files are placed in the 'fhir/add' subdirectory for ingestion. Upon recognition of a new Excel file in this directory, the ETL Tool will automatically trigger the data ingestion process, processing all subjects and their respective clinical data available in the Excel file.

  When processing the files, the data must first be transformed to the correct HL7 FHIR Bundle defined for each cancer with all the information of each patient of each Excel Template.

  Afterwards, each subject's data is individually and promptly posted to the server, with each POST request targeting the server endpoint at http://incisive-dp-fhir:XXXX/baseR4/. Upon receiving these posts, the FHIR Server ingests and manages the incoming data, incorporating it into the datasets within the INCISIVE infrastructure. Through these processes, the ETL tool ensures both the smooth integration and standardization of clinical data, enhancing the overall interoperability of the INCISIVE platform.

Data updating is also streamlined at the patient level. If a patient's information needs to be updated, the updated patient folder simply needs to be placed for ingestion again. The ETL Tool takes over from there, verifying if the patient already exists in the FHIR and/or

PACS systems. If a match is found, the ETL tool will replace all existing data with the new information, thereby ensuring the systems always hold the most current data for each patient.

4.7.5.4. Other complementary tools

The **Quality Check Tool** is developed as a rule-based quality check. Its main purpose is to check whether the data collection requirements are followed and inform the user of potential actions that must be taken to ensure the quality of the data prior to the uploading. Moreover, the tool is extensible, the logic on checking the requirements is not hard-coded, but it is introduced from a knowledge base (specific templates, structures, anonymisation protocol). The tool has 9 components performing checks on clinical metadata, imaging data and the consistency between them.

The **De-Identification tool** is used for de-identifying medical images, specifically DICOM images, by either removing DICOM metadata or modifying them in a secure way. The fields that are modified are based on the DICOM Standard by NEMA, and more specifically DICOM PS 3.15. The tool gives its users the opportunity to select the level of the de-identification they want to apply to their images, by implementing a set of options defined in the aforementioned protocol. Designed and implemented as a web-based application, it can be used by anyone and is also accessible through the Data Sharing Portal of the INCISIVE platform.

The **Transaction Tracker** is a blockchain-based component that is responsible for keeping track the most critical actions performed inside the INCISIVE platform. For each action that is being logged and tracked, a set of information is recorded. This data was defined during meetings with the rest of the consortium and includes information about the username, the kind of the action, the data that was used for this action (if any) and the timestamp of the action. The Tracker also offers an auditing mechanism, through which specific users have the right to check the logs. The Tracker communicates with the platform via a RESTful API, where each service that is exposed corresponds to a different action that is logged. The information that the Tracker receives from the platform, in the form of JSON object and the information that are extracted by the blockchain formulate a JSON object that is then stored in a database. JSON objects are widely used especially when information exchange takes place. The information exchange between the Tracker and the platform takes place using JSON objects, which is a widely adopted format and pairs well with the JavaScript language that the Tracker and its API are implemented with.

4.8. INCISIVE Implementation guide publication

This phase is explained in section 5.3 of this document.

4.9. INCISIVE Recommendations and future actions

This phase is explained in section 6.1 of this document.

5. How to use an existing Interoperability framework or an Implementation guide?

5.1. Roles and interests

These are some of the roles that are involved in the entire process of defining interoperability for a specific use case, following the ReEIF:

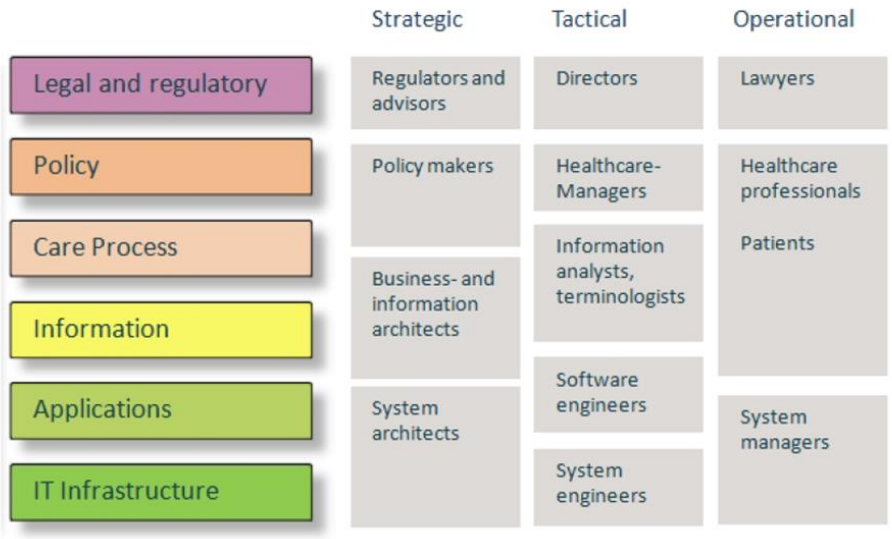


Figure 31 Refined EIF (ReEIF) model – stakeholders.

The Regulators and advisors, the Policy makers, the Directors, the Lawyers will only use the interoperability framework to evaluate the process, the use of data, security, and privacy.

The Business-and information architects, System architects, System managers will use the Interoperability framework, and the implementation guide, to analyse how to implement the designed standard specification.

Software engineers and system engineers will only use the implementation guide, to implement the interoperability platform.

Information analysts and terminologists are the ones who make the standardization suggestions, draft the interoperability framework, and generate the implementation guide.

Healthcare professionals and patients do not use either the interoperability framework or the implementation guide, they only communicate with the Information analysts and terminologists to express their needs regarding the information they would like to view and the data that needs to be obtained.

5.2. Where to look for implementation guides or interoperability frameworks and what to consider when reusing it

Most interoperability standards have a public implementation guide repository. If this is not the case, it is recommended to search directly on the Internet for projects, use cases or assistance processes implementation guides to be sure that what we are going to design, or implement does not already exist.

It is important to consider the following aspects when making the decision to reuse an existing implementation guide:

- The version of the standard with which the implementation guide has been developed.
- The ease with which profiles or extensions can be generated on the guide and standard itself.
- The terminological flexibility to create local codes or personalized.
- The types of data supported or standardized with the proposed standard.
- The ease and standardization proposed by the exchange of clinical documents.

Considering that interoperability frameworks tend to be internal documents, they tend to be reused within the same design team when they must develop more than one interoperability framework for different projects, use cases or processes. Everything that is reusable is kept and the specification is modified for each chosen standard and project.

5.3. How to use the INCISIVE Implementation Guide

The INCISIVE Interoperability Framework is the internal document in the project where it is described all the decisions and explained how need to be implemented the specification for the developer partners of the project, to be sure that everything is reported about the development of the interoperability platform and the standard servers as a service. Also, to have a common framework of all the standardization actions that have been done in INCISIVE.

The INCISIVE HL7 FHIR implementation guide is published on Simplifier.net, that is a FHIR collaboration and publishing platform which can organize all content (e.g., resources and Implementation Guide) in projects, and create, upload, download, find a FHIR resource view. There are a multitude of resources including profiles, Extension, ValueSet, CodeSystem, mappings, examples and more.

The following is an overview of the INCISIVE project on Simplifier[61], which is open and updated in time during the project.

[61] INCISIVE project in SIMPLIFIER.NET. (n.d.). from https://simplifier.net/INCISIVE

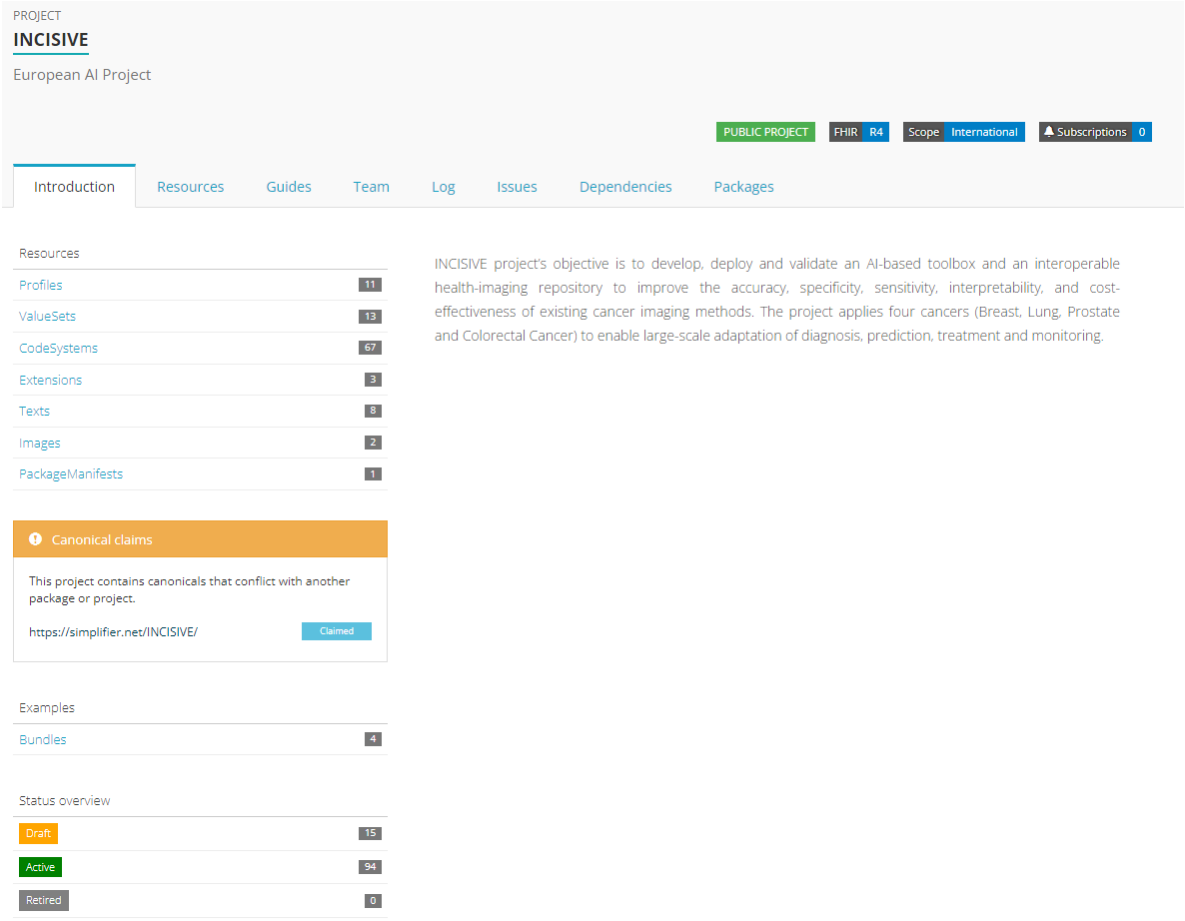


Figure 32 INCISIVE project overview in Simplifier.

The project page contains a menu at the top with different tabs, four tags at the right with basic information, a summary description of the project, and links to resources, examples, and status at the left.

Tags meaning:

- Project privacy: which are whether the project is public or private.
- FHIR version: which version of FHIR is used in this project.
- Scope of project: it can choose International, National, Institute, Regional or Test.
- Number of subscriptions: followers of the project.

Tabs meaning and content:

- Introduction: this section shows an overview of the project.
    - A summary text about the INCISIVE project.

  - A summary table describing the number of resources per different resource type used by the INCISIVE, these numbers of resources will be changed during the project.
  - The number of examples per resource, if enter here, the messages used by the INCISIVE appear there.
  - The canonical base URLs supported in the project are used and recognized as a certificate for resource.
  - The status overview of resource, number of draft resources, active resource and retired resource, these numbers will be changed during the project.
- Resources: this section contains all the Conformance and Example Resources for the INCISIVE HL7 FHIR messages.
- Guides: this section shows Guides for INCISIVE HL7 FHIR messages.
- Dependencies: this section shows the references to other profiles and packages, such as FHIR R4 package that are used by INCISIVE HL7 FHIR messages.
- Packages: this section will show all released packages of the INCISIVE HL7 FHIR messages.

Publishing a project specification in Simplifier is useful for different reasons:

1. Contribute and share a use case specification that can be reused for other projects.
2. Create the URLs for the terminology, extensions, and profiles resources.
3. Create the structure with the simplifier tools and synchronize with forge to be more efficiency and faster designing the resources.
4. Create a complete guide and specification package release in a public and structured environment.
5. Full information for different interested roles, for designers, researchers, implementers, etc.

Below we will explain two search use cases for the INCISIVE project in the simplifier according to an implementer/developer role and a specification interoperability designer role.

5.3.1. For Developers

To start the implementation, it is recommended to download the package of the project in Packages tab that contains all information, which can be immediately installed and used all project's resources.

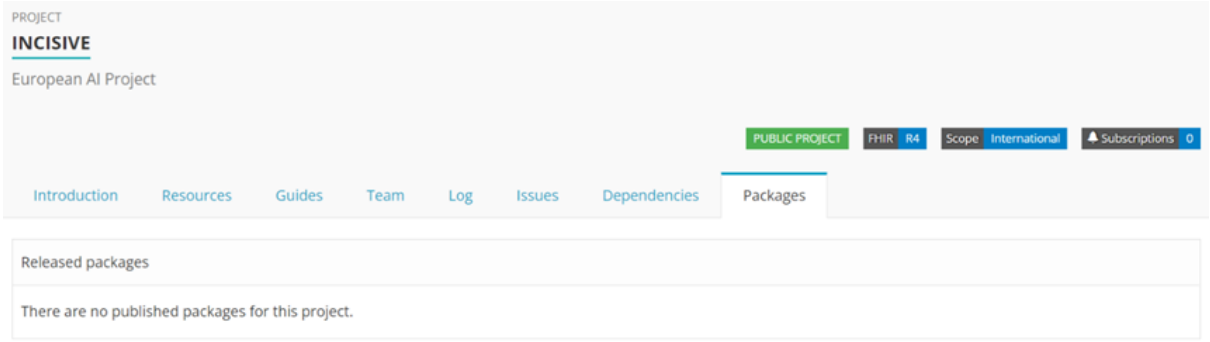


Figure 33 Simplifier Packages view.

With package, the dependencies to other packages can be seen, but they can also be found in the Dependencies tab. Currently, the most important package is FHIR R4, which are definition (API, structures, and terminologies) for the R4 version of the FHIR standards.

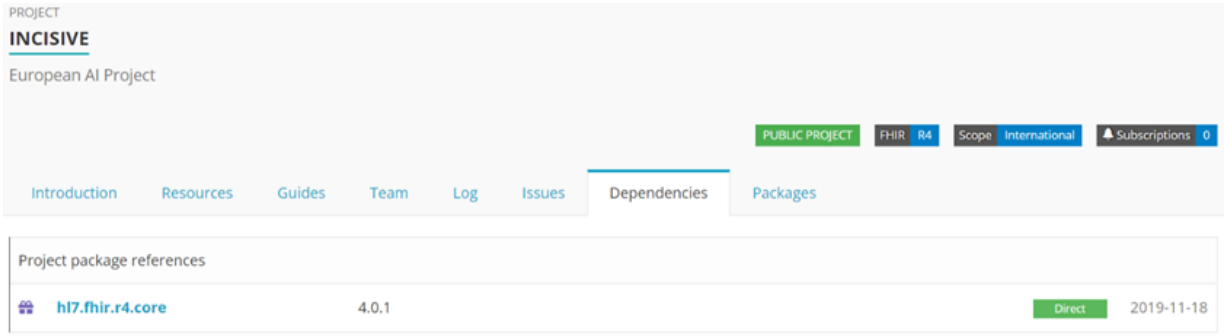


Figure 34 Simplifier Dependencies view.

For interpreting the package, it is useful to enter Guides tab which has a narrative explanation about built project.

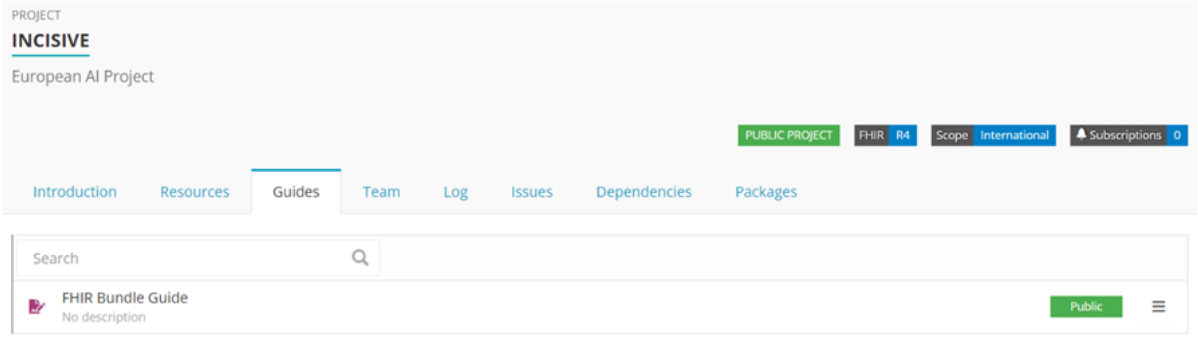


Figure 35 Simplifier Guides view.

On the Home page of the FHIR Bundle Guide, it can be seen the introduction of INCISIVE project and four subpages link in navigation index per cancer type.

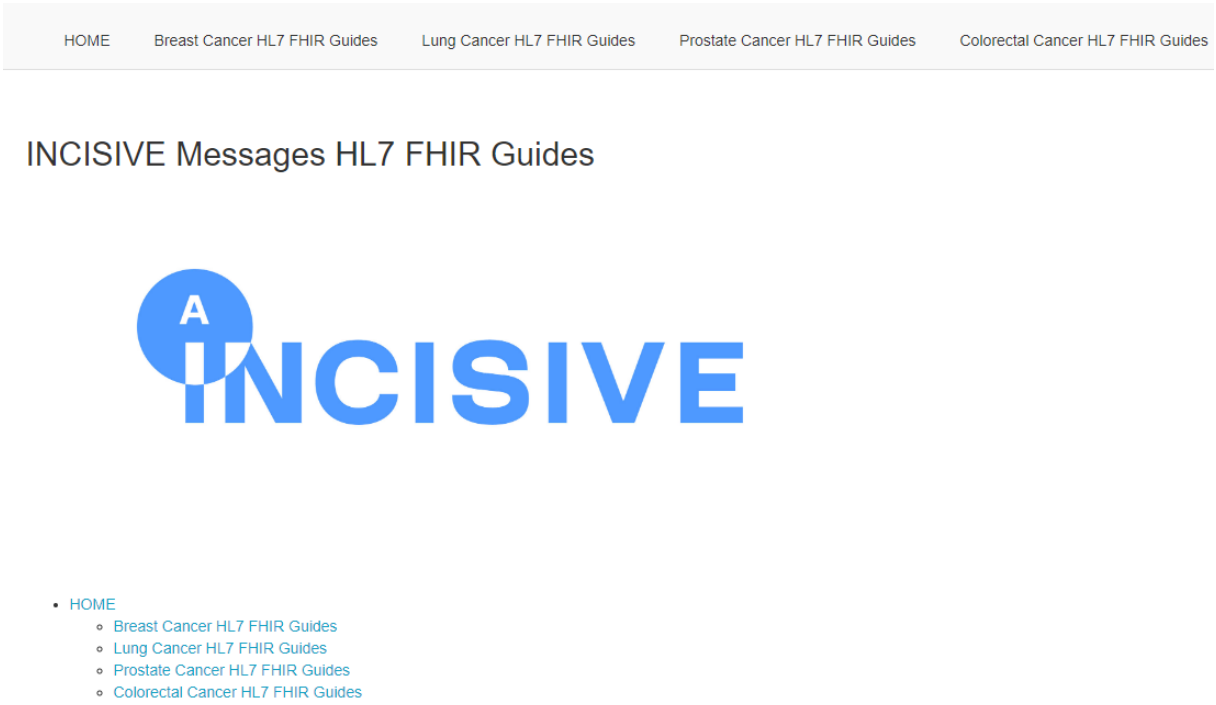


Figure 36 INCISVE bundle guide.

Each subpage shows a guide to the process that needs to be followed to understand the FHIR message we defined.

Breast Cancer HL7 FHIR Guides

1. Introduction

This guide aims to explain the process to follow in order to read and understand the xml message "FHIRbundle_BreastCancer_2023-07-21.xml" based on breast Excel template.

2. Bundle Description

In this chapter describes the bundle content in a table containing these columns.

2.1. Table Columns Description

Column Name	Description
Tab Template	Tab name included in excel file of Breast Cancer
Variable	Variable name included in Excel file
ID resource	ID identification inside the assigned resource
FHIR message location	Attribute used to inform the variable inside the bundle
Example value	Example value used to inform the variable assigned
Required by FHIR specification	Inform the variable is mandatory/optional in XML, based on compliance with the FHIR specification
INCISIVE profile	URL of profiles created by INCISIVE
Condition	Description term condition

2.2. Table with Bundle Description

2.2.1. Tab template **General Info**

Term	ID resource	FHIR message location	Example Value	Required by FHIR specification	INCISVE Profile	Condition
Patient Number	patient01	Patient.identifier.value	004-000001	Optional		
Gender	patient01	Patient.gender	female	Optional		
Ethnicity	patient01	Patient.extension.valueCodeableConcept.coding.code	1	*Optional*	Extension	
		Patient.extension.valueCodeableConcept.coding.display	White	*Optional*		

Figure 37 Breast Cancer bundle guide.

All the dependencies, capability statement and the other resources needs to be supported by the FHIR server that the developer decides to use. On this server some parameters need to be configured and maybe authentication and security services needs to be added.

5.3.2. For Specification Interoperability Designers

It will be very useful to analyse the introduction tab to see the use case and to check if there are similarities with the use case to design.

Then if the use case is similar the Resources tab will be useful to analyse the local codes needed, the terms that have been used, the CodeSystem and semantic design, the profiles and the implementation guide. If there are a lot of similar data with the use case that this wants to design, some parts could be reused or extended.

In this case, it will be useful to download or visualize some examples. The examples are available on the information page at the left side.

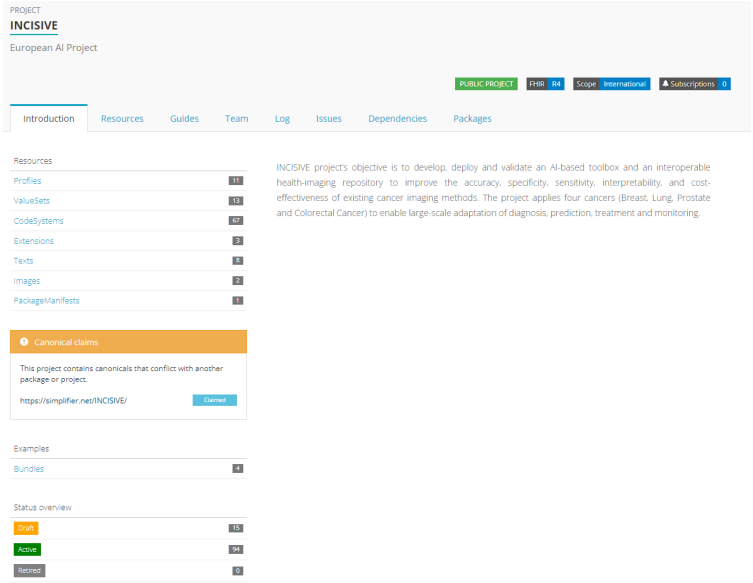


Figure 38 INCISIVE overview in Simplifier.

In the Resource tab, it will find all conformance and terminology resource created for INCISIVE.

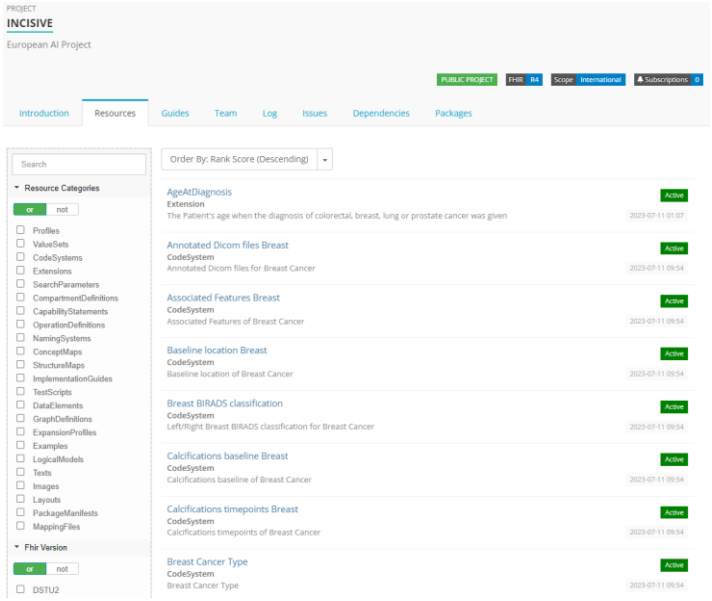


Figure 39 Simplifier Resource view.

When the pages of different types of resources are accessed, they have different overviews.

StructureDefinition resource view: The current interface displays the resource structure, and the diff window only displays the structure designed by INCISIVE instead of the general structure of the resource.

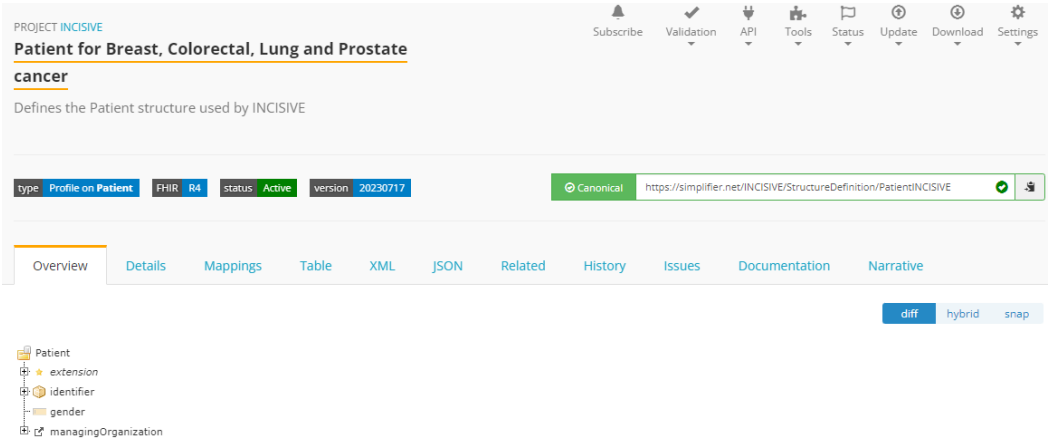


Figure 40 Patient StructureDefinition view.

ValueSet resource view: The current interface displays the CodeSystem contained in the ValueSet, when click the following CodeSystem it will also redirect to the CodeSystem's page.

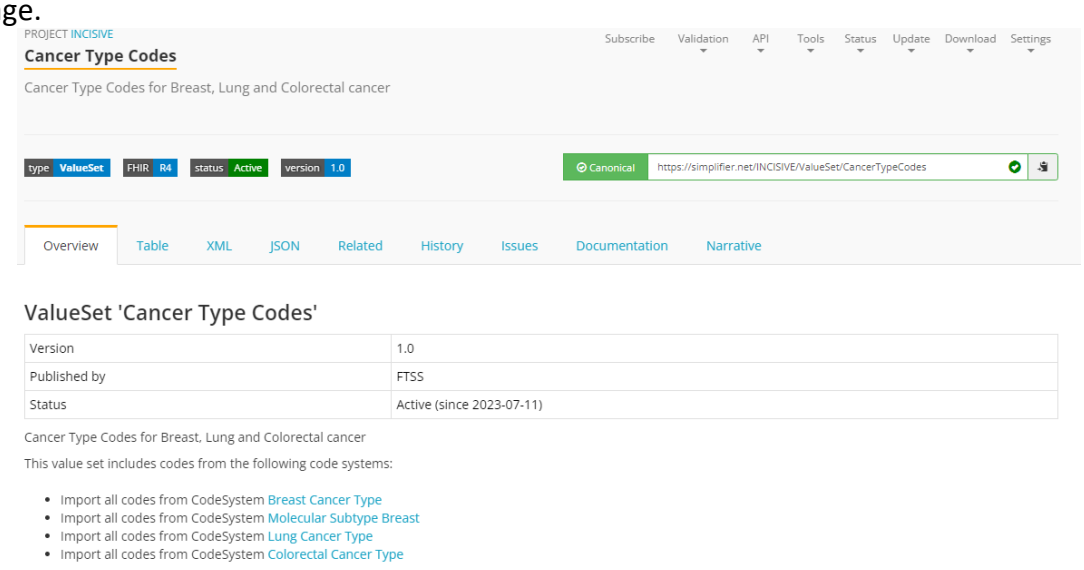
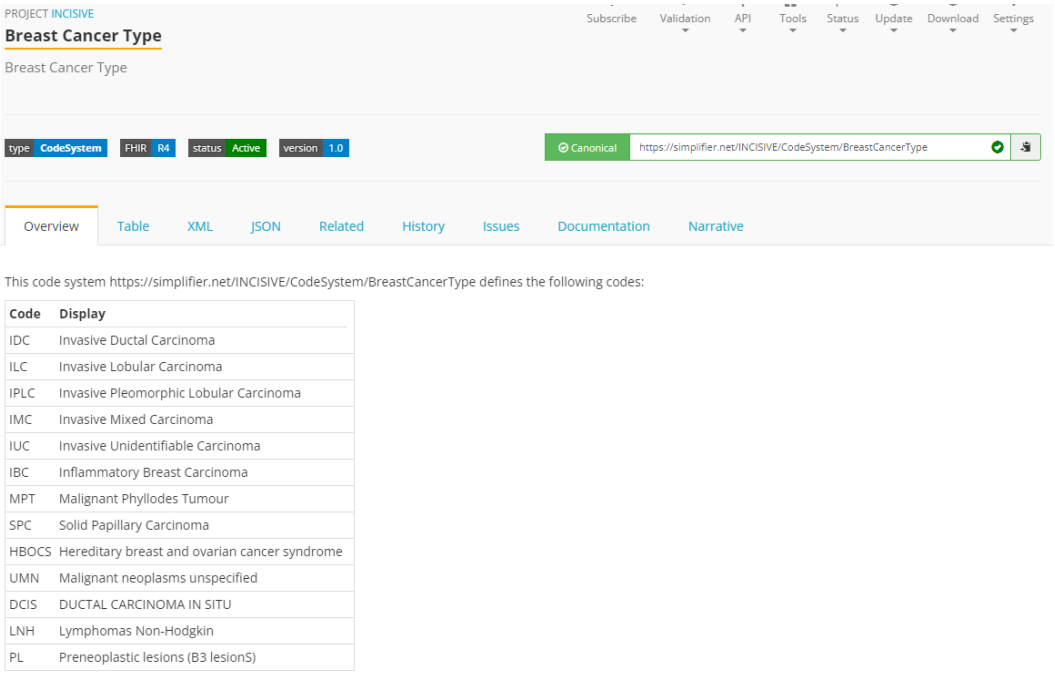


Figure 41 Cancer Type Codes ValueSet view.

CodeSystem resource view: The current interface displays the code list contained in the CodeSystem and the description of the code.



Code	Display
IDC	Invasive Ductal Carcinoma
ILC	Invasive Lobular Carcinoma
IPLC	Invasive Pleomorphic Lobular Carcinoma
IMC	Invasive Mixed Carcinoma
IUC	Invasive Unidentifiable Carcinoma
IBC	Inflammatory Breast Carcinoma
MPT	Malignant Phyllodes Tumour
SPC	Solid Papillary Carcinoma
HBOCS	Hereditary breast and ovarian cancer syndrome
UMN	Malignant neoplasms unspecified
DCIS	DUCTAL CARCINOMA IN SITU
LNH	Lymphomas Non-Hodgkin
PL	Preneoplastic lesions (B3 lesionS)

Figure 42 Breast Cancer Type CodeSystem view.

As you can see, in this role dependencies, package, guides, etc. are not so important. Because the goal is not to implement this specification, the goal is just to reuse, extend, or get inspiration on how to design something similar.

6. Recommendations & Future Actions

This section describes the standardization future actions in INCISIVE, and the general standardization suggestions for future international, European, and local healthcare projects.

6.1. Standardization future actions in INCISIVE

As seen in section 4, INCISIVE has developed its interoperability framework, achieved a common data model for all data providers in the federated platform, in which clinical data is structured in HL7 FHIR messages and medical images in DICOM files, and where all queries are performed via REST APIs to the FHIR Server and PACS. The standardization of the data has made it possible to have a common model to process data from different sources of origin for Model Training in the FNs. In addition, it would also be possible to use OMOP[62] between the Federated Learning Manager (CN) and the Federated Learning Client (FN) when querying the FHIR or PACS server to index data either with query profiles, with cases of use with different visions according to the type of cancer, etc.

Regarding data uploading use cases, INCISIVE receives data for Model Training and Model Inference. For Model Training to facilitate joining as a data provider, the technical impact for providers has been minimized, so the option to upload an Excel file for clinical data, and upload DICOM files for medical images, avoiding the need for integration and own development of a FHIR API or a REST API for each data provider. At a production stage, the ETL tool should be dispensed with as much as possible, data providers should directly POST HL7 FHIR messages for clinical data and should POST medical images directly to the Orthanc PACS. In short, the ETL tool should be replaced by a WebService to publish data to the FHIR Server and Orthanc PACS by all data providers using the appropriate API. For Model inference, to facilitate the data loading process for healthcare professionals, an Excel file for clinical data and DICOM files for medical images have been chosen as the only method without an ETL tool. In the production stage, the model inference loading process should also use a FHIR API and a PACS REST API and should have data persistence, since the same information could be reused for the Models Training, as retraining.

Regarding the FHIR server, INCISIVE has chosen to use an IBM FHIR server because it is a java open-source solution and from an internationally recognized organization. But there are different FHIR servers available, both open source and proprietary solutions for different major cloud providers. The FHIR server provides us with different functions such as data storage, Rest API, profile validation and user interface for interaction. Different FHIR servers are available in the market including open-source solutions like HAPI and IBM FHIR servers or proprietary solutions

[62] OMOP Overview. (n.d.). from https://www.ohdsi.org/data-standardization/, https://www.ohdsi.org/2019-tutorials-omop-common-data-model-and-standardized-vocabularies/, https://ohdsi.github.io/CommonDataModel/

from different cloud providers GCP(Google), Azure, AWS etc. The HAPI FHIR server[63] is the most famous of the open-source FHIR servers that are available. HAPI is an open community that develops software licensed under the Apache Software License 2.0. HAPI FHIR is a product of Smile CDR. You have a docker desktop or frontend to access the server at http://localhost:8080. For all these reasons it is recommended to use an open-source FHIR server, and it is recommended to use the HAPI FHIR server because it is the most used and has this front access which for research projects could be very useful to make demos about the data exchange.

Regarding HL7 FHIR queries, it is recommended to review all parameters that API FHIR has and see how it can take full advantage of all its benefits, previously configured in the CapabilityStatement.

Regarding the HL7 FHIR specification, in INCISIVE is used the R4 version because it is the normative version of FHIR. But it is very important to analyse and start thinking about the R5 version which will be the next version. In INCISIVE a study of this update was done, and only some resources will need to be updated:

- No changes for FHIR Resources: Patient, MessageHeader and Bundle.
- Changes that will not affect INCISIVE FHIR Resources: Observaton, DiagosticReport, Organization, Medication.
- Changes that will require an update for INCISIVE FHIR Resources: Condition.evidence mandatory, Procedure.performedString.usedCode to Procedure.occurrenceString, MedicationAdministration.medicationReference to MedicationAdministration.medication, Delete MedictionStatement.
- Upgrade the IBM FHIR server to R5 version, if supported.

Regarding the HL7 CDA Clinical Report for the results of the AI Services of the Model Inference, a direct integration with the information systems of the healthcare professional is recommended for the exchange of this document and the subsequent exploitation or register. There are also some CDA attributes that are not included in the XML version available for download, which are recommended to be added in production, and they are the following optional attributes:

- More information about the Clinical Report Patient and Healthcare professional:

recordTarget.patientRole.patient.name.given	Is recommended to know this information of the Clinical Report Patient. But if the Patient information is anonymised, it is only informed the patientID.
recordTarget.patientRole.patient.name.familiy	
recordTarget.patientRole.patient.administrativeGenderCode	
recordTarget.patientRole.patient.birthTime	
recordTarget.patientRole.providerOrganization.id	
Author.assignedAuthor.assignedPerson.name.given	Is recommended to know this information of the Clinical Report Healthcare professional (HP). But if the
Author.assignedAuthor.assignedPerson.name.family	

[63] HAPI FHIR. (n.d.). from http://hapi.fhir.org/

Author.assignedAuthor.assignedPerson.name.suffix	HP information is anonymised, it is only informed the hpID.

Table 16 Clinical Report Patient and HCP attribute.

- More information about the Clinical Report Legal Authenticator:

legalAuthenticator.time	Is recommended to represents a participant who legally authenticate the report.
legalAuthenticator.signatureCode	
legalAuthenticator.assignedEntity.id	
legalAuthenticator.assignedEntity.assignedPerson.name.given	
legalAuthenticator.assignedEntity.assignedPerson.name.family	
legalAuthenticator.assignedEntity.assignedPerson.name.suffix	
legalAuthenticator.assignedEntity.representedOrganization.id	

Table 17 Clinical report Legal Authenticator attribute.

- More information about the Clinical Report Life Cycle:

relatedDocument.parentDocument.id	Is recommended to indicate the revision, modification, and transformation of the report.
relatedDocument.parentDocument.setId	
relatedDocument.parentDocument.versionNumber	

Table 18 Clinical Report Life Cycle attribute.

- More information about the Clinical Report Encounter:

componentOf.encompassingEncounter.id	Is recommended to indicate the setting of the clinical encounter during which documentation is carried out.
componentOf.encompassingEncounter.effectiveTime	
componentOf.encompassingEncounter.encounterParticipant.time	
componentOf.encompassingEncounter.encounterParticipant.assignedEntity.id	
componentOf.encompassingEncounter.encounterParticipant.assignedEntity.assignedPerson.name.given	
componentOf.encompassingEncounter.encounterParticipant.assignedEntity.assignedPerson.name.family	
componentOf.encompassingEncounter.encounterParticipant.assignedEntity.assignedPerson.name.suffix	
componentOf.encompassingEncounter.encounterParticipant.assignedEntity.representedOrganization.id	
componentOf.encompassingEncounter.location.healthCareFacility.code	

Table 19 Clinical Report Encounter attribute.

It is also recommended to include this HL7 CDA message in an HL7 FHIR message and add the FHIR document exchange method, considering also the IHE XDS recommendations.

6.2. General standardization suggestions

As seen in this deliverable and throughout the INCISIVE standardization process, interoperability is a necessity for any system that needs to exchange data with another system. In the health sector information systems must be able to communicate with each other both at the national

level, and/or at the European level, and/or at the international level. At the same time, it is necessary to standardize not only the data, but the methodology with which the interoperability solutions are proposed and with which the standards to be used for each use case are decided.

In terms of standards, the key is to use the most appropriate standard for each functionality and combine them with each other to achieve the most complete standardized service. It seems that looking for a standard that allows us to standardize everything is not the solution that will best fit all our data, but the standards that will be used the most will be those that can complement others and that can be adapted to local needs or specific processes through profiles and extensions.

Considering all this, the specification proposed by INCISIVE combines the standards in the most suitable way for the data it handles, SNOMED CT and LOINC are the most used terminologies to encode clinical and laboratory data, DICOM is the most used global standard for medical imaging, HL7 FHIR is the most used standard internationally for the exchange of clinical data encoded, HL7 CDA is the most used global standard for clinical documents and IHE is the most used global standard for the definition of clinical use cases and processes.

However, it should be noted that the interoperability platform could still be improved, considering the following aspects that are emerging in the market:

6.2.1. Regarding use cases and healthcare process definition

When starting to define an interoperability scenario it is important to consider the context, domain, existing systems and future needs of each use case or process. But often these use cases or care processes are common in many places in the world. For this reason, there are organizations like IHE that define these processes and their possible solutions to standardize them globally.

Integrating the Healthcare Enterprise (IHE) is an initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information. It promotes the coordinated use of established standards such as DICOM and HL7 to address specific clinical needs in support of optimal patient care and guarantees better system communication, easier implementation and more effectively to enable care providers to use information. Is organized by clinical and operational domains. In each domain users with clinical and operational experience identify integration and information sharing priorities and vendors of relevant information systems develop consensus, standards-based solutions to address them. Each domain develops and maintains its own set of Technical Framework documents.

These are the IHE domains: Cardiology, Dental, Devices, Endoscopy, Eye Care, IT Infrastructure, Pathology and Laboratory Medicine, Patient Care Coordination, Pharmacy, Quality and Research and Public Health, Radiation Oncology, Radiology.

The IT Infrastructure domain is the one that includes the "Actor Groupings" which are referenced by other domains, and which explain how combined they can offer very complete functionalities for the health sector.

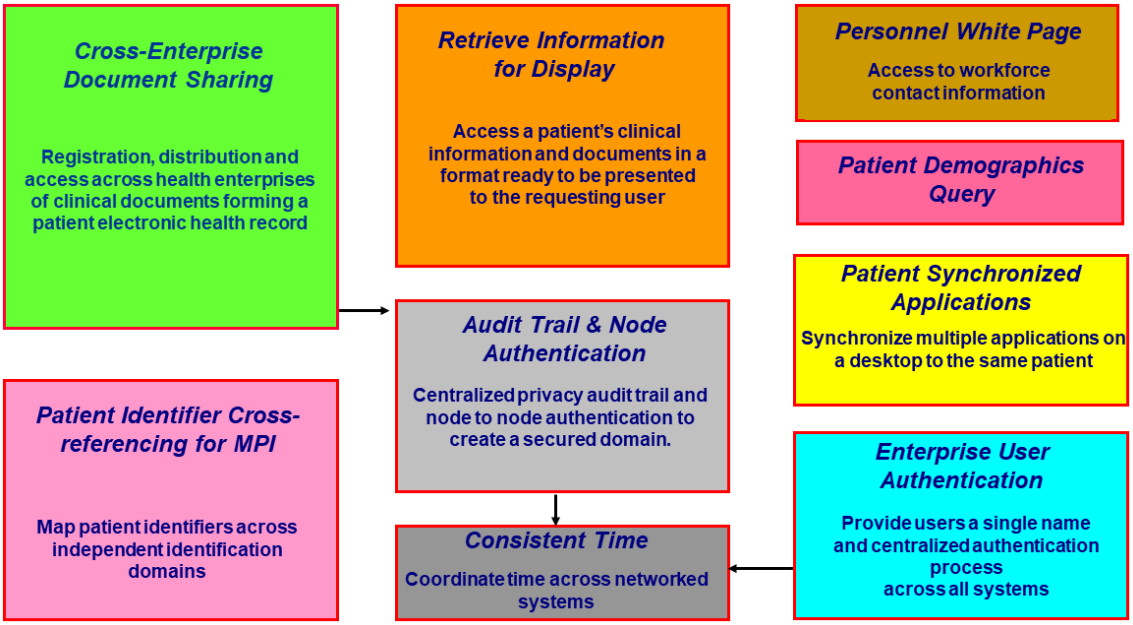


Figure 43 Actor Groupings in health sector.

In addition, auditing in any health information system is often mandatory and IHE also defines technical audit or security implementation profiles such as the Audit Trail and Node Authentication (ATNA) profile[64] that specifies the foundational elements needed by all forms of secure systems: node authentication, user authentication, event logging (audit), and telecommunications encryption. It is also used to indicate that other internal security properties such as access control, configuration control, and privilege restrictions are provided.

In conclusion, considering the IHE domains when defining use cases for any process or project in the health sector is a key piece to a good definition and implementation of interoperability.

Furthermore, ISOs, EMDN and MDC, OASIS are other international standards mandatory to consider for solutions especially in healthcare sector.

In sections 3.2.7.1, 4.7.1 and 4.7.1.1 different International Organization for Standardization (ISO), Medical Device Certification (MDC) etc. have been discussed.

[64] ATNA IHE Overview. (n.d.). from https://wiki.ihe.net/index.php/Audit_Trail_and_Node_Authentication

Regarding healthcare devices regulations the Medical Device Regulation (MDR) and In Vitro diagnostic medical Devices Regulation (IVDR), the European Medical Device Nomenclature (EMDN) aims at supporting the functioning of the European Database on Medical Devices (EUDAMED). Among its various uses, it will be utilised by manufacturers for the registration of medical devices in EUDAMED, where every medical device will be associated to each Unique Device Identifier – Device Identifier (UDI-DI). It also plays a key role in MDR/IVDR device documentation and technical documentation, sampling of technical documentation conducted by notified bodies, post-market surveillance, vigilance, and post-market data analysis, etc. It is intended to support all actors in their activities under the MDR/IVDR and provides key device descriptions to patients as regards their own devices and all other devices available on the market and registered in EUDAMED.

Furthermore, Organization for the Advancement of Structured Information Standards (OASIS) standard[65] provides guidelines about best practices in writing specifications, so that the risk of having interoperability (or portability) failures between implementations is reduced.

6.2.2. Regarding the exchange of data without losing the semantic meaning

Preserving semantic meaning is crucial for maintaining the accuracy and integrity of clinical information.

When humans and systems use different words and codes to describe the same thing, dependent data types and values can generate different medical concepts, if the data is inaccurate or misunderstood, it can cause medical errors. Therefore, there is a need to use terminologies to represent the meaning of data and help maintain medical concepts and information.

There are various standardized clinical terminologies that can be used depending on the type or use of data, such as SNOMED CT, LOINC, RxNorm, ICD, NANDA, UCUM, CPT, MED-RT, NDF-RT, UNII, NDC, CVX, HPO, MEDCIN, ATC, etc. These terminologies provide a common vocabulary that allow all members of a healthcare system to understand the data meaning in the same way; independently of the methods or standards selected for the exchange data and independently of the language, ensuring consistent interpretation and meaning across different healthcare systems.

[65] *OASIS Overview. (n.d.). from https://www.oasis-open.org/policies-guidelines/interoperability-guidelines/*

Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT)[66] is the terminology used to code, retrieve, communicate, and analyse clinical data, enabling healthcare professionals to express clinical data accurately and unambiguously.

Logical Observation Identifiers Names and Codes (LOINC)[67] is the terminology used to identify laboratory data such as laboratory tests, measurements, etc.

RxNorm[68] provides normalized names for clinical drugs and links its names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software, including those of First Databank, Micromedex, Multum, and Gold Standard Drug Database.

International Statistical Classification of Diseases and Related Health Problems (ICD)[69] is the terminology used to code extent, causes and consequences of human disease and death worldwide.

NANDA International's Nursing Diagnoses[70] is the definitive guide to nursing diagnoses, providing the critical information needed for nurses to understand assessment, its link to diagnosis and clinical reasoning, and the purpose and use of taxonomic structure for the nurse at the bedside.

Unified Code for Units of Measure (UCUM)[71] is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business.

Current Procedural Terminology (CPT)[72] is the terminology offer doctors and health care professionals a uniform language for coding medical services and procedures to streamline reporting, increase accuracy and efficiency.

National Drug File Reference Terminology (NDF-RT)[73] is the terminology used for modelling drug characteristics including ingredients, chemical structure, dose form, physiologic effect, mechanism of action, pharmacokinetics, and related diseases.

[66] SNOMED CT Overview. (n.d.). from https://www.snomed.org/
[67] LOINC Overview. (n.d.). from https://loinc.org/
[68] RxNorm. (n.d.). from https://www.nlm.nih.gov/research/umls/rxnorm/index.html
[69] ICD Overview. (n.d.). from https://www.who.int/standards/classifications/classification-of-diseases
[70] NANDA Overview. (n.d.). from https://nanda.org/
[71] UCUM Overview. (n.d.). from https://ucum.org/
[72] CPT Overview. Dotson, P. CPT® Codes: What Are They, Why Are They Necessary, and How Are They Developed? Advances in Wound Care. 2013 Dec; 2(10) from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3865623/
[73] NDF-RT Overview. (n.d.). from https://evs.nci.nih.gov/ftp1/NDF-RT/NDF-RT%20Documentation.pdf

Medication Reference Terminology (MED-RT)[74] is the successor to the Veterans Health Administration National Drug File Reference Terminology (VHA NDF-RT™). It is formal ontological representations of medication terminology, pharmacologic classifications, and asserted relationships between them.

Unique ingredient identifier (UNII)[75] is an alphanumeric identifier linked to a substance's molecular structure or descriptive information and is generated by the Global Substance Registration System (GSRS) of the Food and Drug Administration (FDA).

FDA's National Drug Code (NDC)[76] Directory contains information about finished drug products, unfinished drugs, and compounded drug products.

Vaccine administered code set (CVX)[77] is a HL7 standard code set. CVX codes for inactive vaccines allow transmission of historical immunization records. These codes should be used for immunization messages using either HL7 Version 2.3.1 or HL7 Version 2.5.1.

The Human Phenotype Ontology (HPO)[78] provides a standardized vocabulary of phenotypic abnormalities encountered in human disease.

The Anatomical Therapeutic Chemical (ATC)[79] Classification is an internationally accepted classification system for medicines that is maintained by the World Health Organisation (WHO) that assigns ATC codes to all active substances contained in medicines based on the therapeutic indication for the medicine.

MEDCIN[80] is a medical terminology that encompasses symptoms, history, physical examination, tests, diagnoses, and therapies.

In summary, there are different specialized terminologies for different medical concepts, and it is often necessary to agree on them to code each term with the terminology that best suits it. Usually, all these terminologies can be referenced within the structures marked by data exchange standards.

[74] MED-RT Overview. (n.d.). from https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/MED-RT/index.html
[75] UNII Overview. (n.d.). from https://www.fda.gov/industry/structured-product-labeling-resources/uniis-preferred-substance-names-and-their-identified-synonyms
[76] NDC Overview. (n.d.). from https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory
[77] CVX Overview. (n.d.). from https://www2a.cdc.gov/vaccines/iis/iisstandards/vaccines.asp?rpt=cvx, https://www.nlm.nih.gov/research/umls/rxnorm/sourcereleasedocs/cvx.html
[78] Human Phenotype Ontology Overview. (n.d.). from https://hpo.jax.org
[79] ATC Overview. (n.d.). from https://www.who.int/tools/atc-ddd-toolkit/atc-classification
[80] MEDCIN Overview. (n.d.). from https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/MEDCIN/index.html

6.2.3. Regarding Clinical Documents exchange

The exchange of documents using formats such as .docx or .pdf makes it difficult to post-interpret and process the data between different information systems. This is why it is necessary to make use of a document exchange standard.

Many information systems around the world have adopted the HL7 v2 or v3 standard as a messaging standard and then the HL7 CDA as a clinical document exchange standard. Currently, it seems that the messaging standard for the exchange of clinical data between information systems of the future is HL7 FHIR, and many states are already adopting this standard. Both IHE and HL7 are working on an updated proposal for the exchange of clinical documents and make proposals such as FHIR Documents or IHE XDS, compatible with each other.

HL7 Clinical Document Architecture (CDA) is an XML-based standard for capturing and exchanging clinical documents, such as discharge summaries or progress notes. It defines the structure and semantics of the documents to ensure interoperability. CDA utilizes standardized clinical terminologies and enables the inclusion of both structured and unstructured data.

HL7 FHIR enables the exchange of clinical documents using the data exchange method with the Documents Module. A FHIR document instance is a Bundle of type document that starts with a Composition resource and contains specific frozen versions of other resources and specific rules based on the type of document (e.g. legal documents, clinical documents, etc.). Documents built in this module may be exchanged between systems and persisted in document storage and management systems, including systems such as IHE XDS. HL7 has plans to develop profiles in the future giving additional guidance on appropriate representation of clinical documents in general, specific types of clinical documents (e.g. Consolidated CDA), and other non-clinical documents (Death Certificate, Medication Registration Information). Note that FHIR defines both this document format and a DocumentReference resource. FHIR documents are for documents that are authored and assembled in FHIR, while the DocumentReference resource is for general references to documents (which may include FHIR documents as well as PDFs, CDAs, etc.).

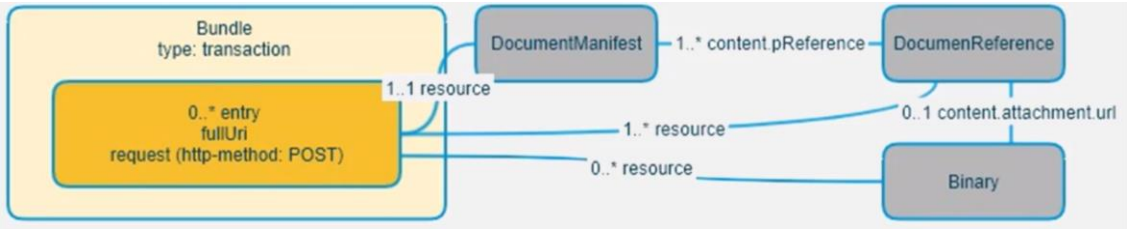


Figure 44 Example of HL7 FHIR exchange method.

IHE Cross Enterprise Document Sharing (XDS)[81] is a system of standards for cataloguing and sharing patient records across health institutions. XDS provides a registry for querying which patient records are in an EHR repository and methods for retrieving the documents. The XDS system of registry and repository is termed an integration profile and was created by Integrating the Healthcare Enterprise (IHE). An XDS document is the smallest unit of information that can be provided to a document repository and registered as an entry in the document registry actor. Also, is a composition of clinical information that contains observations and services for the purpose of exchange with the following characteristics: Persistence, Custody, Authentication Potential, and Completeness. These features are defined in the HL7 Clinical Document Architecture specification. It must comply with a published standard that defines its structure, content and encoding.
IHE defines content-oriented integration profiles based on these content standards for use in conjunction with XDS. The XDS integration profile manages XDS documents as a single unit of information; does not provide mechanisms for accessing parts of an XDS document. Only document sources or document consumers have access to the internal information of the XDS document. Therefore, we could say that IHE XDS is an architectural approach that separates document management from metadata management.

[81] IHE Cross Enterprise Document Sharing Overview. (n.d.). from https://wiki.ihe.net/index.php/Cross-Enterprise_Document_Sharing

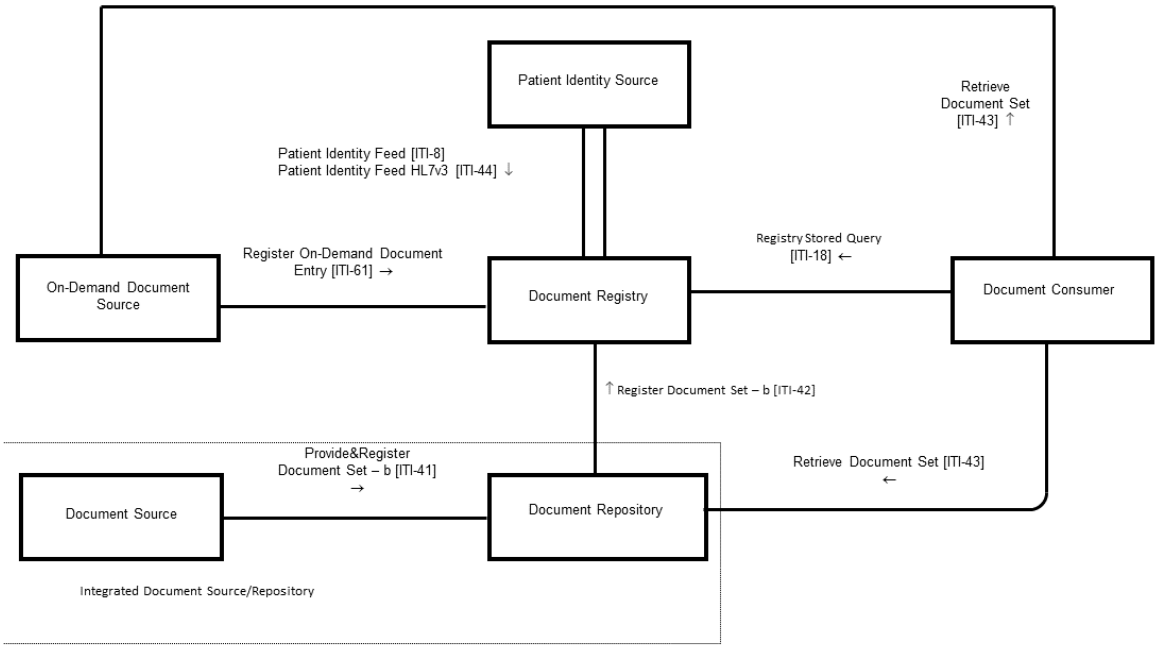


Figure 45 XDS structure.

In summary, HL7 and IHE are establishing a document exchange method. IHE XDS appears to maintain the use of HL7 CDA and provides a use-case implementation architecture model across different systems. FHIR Documents seem to replace HL7 CDA in the sense that coming from the same family and taking into account the advantages of FHIR over HL7 v2 or v3 also apply to this proposal achieving with FHIR resources to report the same as with an HL7 CDA, also allow to attach the pdf, image, resource and at the same time solve the architecture proposed by IHE with the FHIR API.

Therefore, when deciding to use one approach or another, it seems reasonable to see what standards are being used by the different information systems that will act as sources and consumers, and to see if it is necessary to be more conservative in applying IHE XDS that basically be to add the transactions and actors that are missing in our current process; or opt for a new proposal such as FHIR Documents, especially if some of the sources or consumers have FHIR servers, in order to enjoy the advantages that FHIR brings.

Also, it seems that IHE XDS and HL7 FHIR are finding synergies to coexist, to map and complement each other. Although it should be noted that IHE XDS attributes do not have the same cardinality as HL7 FHIR, and this makes FHIR less restrictive in this regard, when considering a coexistence between these two standards for interchange of documents, the cardinalities indicated by IHE XDS must be maintained on FHIR. At the same time, it seems easier to query FHIR resources using the FHIR API, than with the implementation

proposed by IHE XDS. In conclusion, it is possible to follow the IHE XDS model by making use of the HL7 FHIR Document Module and its API benefits.

6.2.4. Regarding Medical Image exchange

The exchange of medical images using formats such as .jpg or .png makes it difficult relate to patient information and to interpret or process between different information systems. This is why it is necessary to make use of a medical image exchange standard.

The most used standard by far is Digital Imaging and Communications in Medicine (DICOM) which handle, store, print, and transmit medical imaging and related information. It includes a file format definition and a network communications protocol. DICOM files can be exchanged between two entities that can receive image and patient information in DICOM format. DICOM was developed to allow the integration and communication of scanners, radiographic devices, servers, workstations, and multiple hardware, from different suppliers.

It is common to use the Picture Archiving and Communication System (PACS) to manage and store the DICOM images and to enables share medical images within their organizations and ensures that healthcare professionals will be able to get a chronological measure of the imaging history of their patients. Also, DICOMLibrary has a PACS storage calculator available for free to know the storage size to use for any study type and use case.

DICOM includes various service classes, object formats, and communication protocols to facilitate image exchange, such as: DICOMweb, DICOM SOP, DICOM SR, etc.

DICOMweb[82] is a set of RESTful services, that can be implemented directly or as a proxy to the DIMSE services to offer modern web-based access to DICOM-enabled systems.

DICOM Service-Object Pair (SOP)[83] is a service defined in a class by the union of an Information Object Definition (IOD), e.g. CT images, MR images, but also include schedule lists, print queues, etc.; and a DICOM Service Elements (DIMSE), e.g. Store, Get, Find, Move, etc. The SOP Class definition contains the rules and semantics which may restrict the use of the services in the DIMSE Service Group or the Attributes of the IOD.

DICOM Structured Report (SR)[84] is the diagnostic report that encodes the interpretation and the impressions of the radiologist about the DICOM image in a structured header and a document body.

[82] DICOMweb (n.d.). from https://www.dicomstandard.org/using/dicomweb
[83] SOP Overview. (n.d.). from https://www.dicomlibrary.com/dicom/sop/
[84] SR Overview. (n.d.). from https://dicom.nema.org/dicom/2013/output/chtml/part20/sect_A.3.html

It is important to use and know all these services and complement them with other data exchange standards like HL7 v2 or v3, IHE XDS, HL7 FHIR, IHE Radiology domains, to address specific use cases or workflow scenarios, such as image sharing, radiology reporting, or patient identification.

6.2.5. Regarding Clinical Data and Events exchange

It has been explained in sections above the standardizing use cases and processes, maintaining the meaning of clinical terms through terminologies, standardizing the structure of documents, using a standard that allows relating patient information to a medical image standard. But all this is somehow sent between one actor and another or between one entity/system and another, and this method of exchanging clinical data also needs to be defined in a standardized way.

The most used standard for the exchange of clinical data is Health Level 7 (HL7) version 2 or version 3 which is an internationally recognized standard for exchanging clinical and administrative data. It provides a framework for structuring, encoding, and transmitting healthcare information electronically.

The evolution of the HL7 v2 or v3 standard is Fast Healthcare Interoperability Resources (FHIR), which is also an HL7 standard. More than evolution, it is a rethinking with many more facilities, maintaining the same functionalities and adding more. At the same time allowing it to coexist and be mapped with the HL7 v2 and v3 and other standards.

HL7 FHIR utilizes a resource-based approach and employs standard healthcare terminologies. It focuses on simplicity, flexibility, and extensibility, allowing for the exchange of granular clinical data. Also utilizes a RESTful API, which aligns well with modern web-based technologies and enables easy integration with existing systems. This approach leverages widely adopted web standards, such as HTTP and JSON, XML, making it simpler to develop and integrate FHIRbased applications. It allows for modular implementations, enabling organizations to adopt and implement specific resources or profiles relevant to their needs. This flexibility allows for incremental adoption, making it easier to integrate it into existing healthcare systems.

Furthermore, it has a thriving ecosystem with a large community of developers, implementers, and vendors actively contributing to its development and adoption. This means that there are abundant resources, tools, and libraries available for implementing FHIR-based solutions.

In summary, FHIR provides the following benefits to developers of ICT solutions in health systems: data exchange interfaces (APIs), accessibility to multiple open-source online tools including reference servers and libraries, solving the needs of information generators and consumers, to share data in a "light-weight" way and in real time, using

the just-in-time interoperability with basic resources that can be used out-of-the-box, but can also be adapted to address local requirements by creating specific profiles, coexistence and synergy with the classic HL7 v2 o v3, HL7 CDA and other standards, use of web standards such as REST, XML, JSON, HTTP and OAuth, online specification updated and easy to understand, a human-readable data serialization format that makes it easy for developers to use, alignment with the technical frameworks and integration profiles of IHE, mappings with different standards for example OMOP CDM or OpenEHR.

Finally, considering data mapping, transformation, validation, and quality is critical to data exchange between systems. Data mapping and transformation techniques can be employed when exchanging information between systems that use different formats or terminologies. This involves mapping the source data elements to the corresponding elements in the target system, preserving the semantic meaning during the translation process. UMLS is the best tooling resources, mapping more than two hundred clinical vocabularies. Implementing data validation and quality checks is crucial to ensure the accuracy and integrity of exchanged clinical data. These checks can include validating data against defined standards, performing semantic validation to ensure consistency, correctness and verifying data integrity during the exchange process.

6.2.6. Regarding CDM to index or group data

Many times, DBB[85] must be queried with different views (roles, domains, usage) and give fast response to many requests per minute, coming from various target applications. In this case, you can use APIs to consult DBB such as ElasticSearch[86], Algolia[87] or use some standard as an analysis layer.

The Observational Medical Outcomes Partnership (OMOP)[88] was a public-private partnership involving the FDA, multiple pharmaceutical companies, and healthcare providers established to inform the appropriate use of observational healthcare databases for studying the effects (risks and benefits) of medical products. It provides a structured representation of healthcare data, including some clinical, demographic, and administrative information.

[85] Databases
[86] Elasticsearch Platform Overview(n.d.). from https://www.elastic.co/
[87] Algolia Overview. (n.d.). from https://www.algolia.com/
[88] OMOP Overview. (n.d.). from https://www.ohdsi.org/data-standardization/, https://www.ohdsi.org/2019-tutorials-omop-common-data-model-and-standardized-vocabularies/, https://ohdsi.github.io/CommonDataModel/

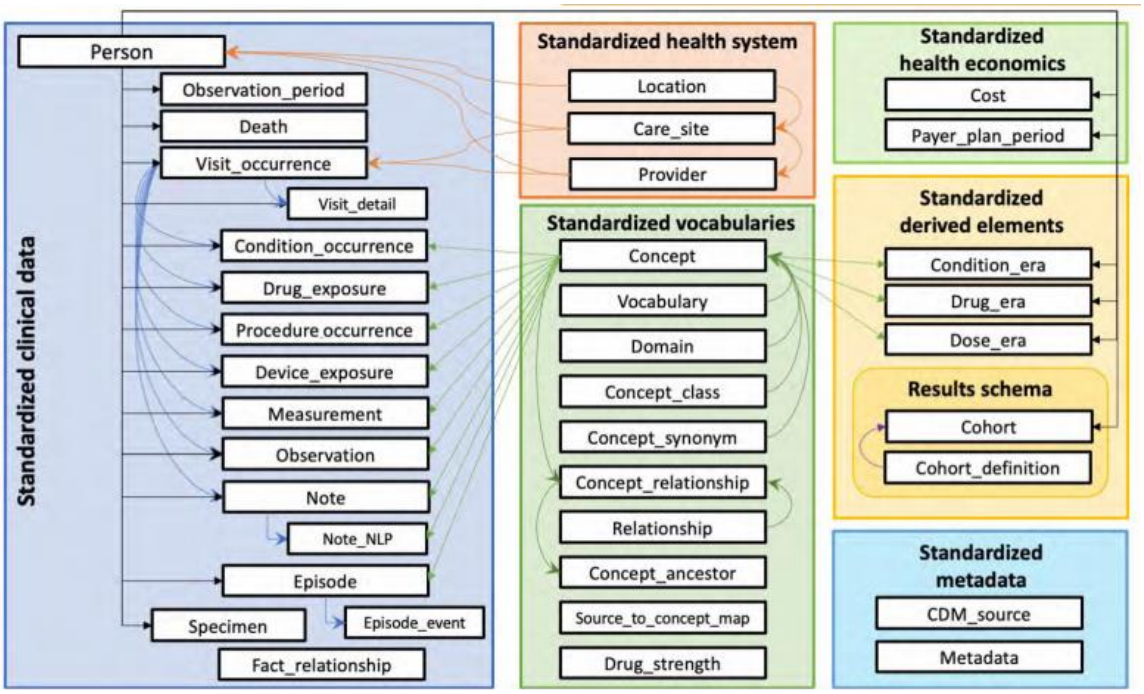


Figure 46 Common data model.

The OMOP Common Data Model (CDM) is an open community data standard, designed to standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence. OMOP was specifically designed to support observational research studies and evidence generation. A central component of the OMOP CDM is the OHDSI standardized vocabularies. The Observational Health Data Science and Informatics (OHDSI) vocabularies allow organization and standardization of medical terms to be used across the various clinical domains of the OMOP common data model and enable standardized analytics that leverage the knowledge base when constructing exposure and outcome phenotypes and other features within characterization, population-level effect estimation, and patient-level prediction studies.

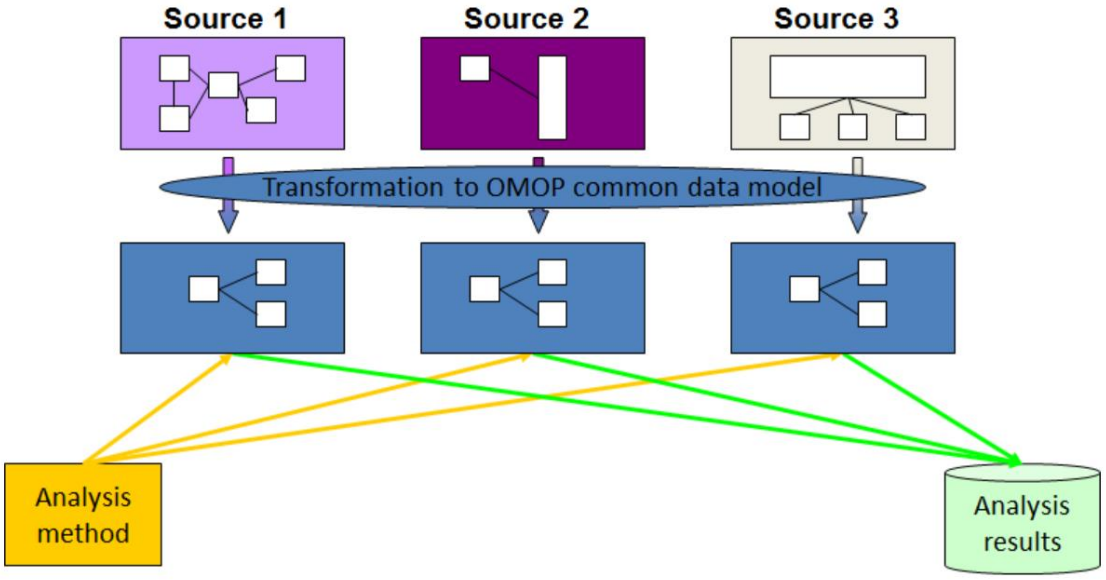


Figure 47 Common data model structure.

OMOP is a relatively new standard, although it is already being used in many information systems and appears to be on the growth curve. However, many implementers comment that transforming the entire current domain of an information system into OMOP often requires the use of external extensions or the use of other standards that supplement it with more domains to cover all the information and events that occur in an information system.

The CDM proposed by OMOP is indeed very powerful, especially for data analysis, as a consumer of this data. In other words, OMOP as a messaging standard might not be the most suitable, since it requires many extensions to be able to cover all the domains of a service with different use cases, but yes for the consumption of data, hosted to different FHIR or non-standardized DB servers, which require to be analysed or used together. For example, it is a recommended standard for indexing standardized data, making views of the data of a DB, defining profiles of this data according to query user, etc.

Also, by incorporating medical images (associated metadata, such as patient demographics, imaging modality, acquisition parameters, and radiology reports) into the CDM, researchers can perform analyses that combine imaging data with other clinical data. This could lead to new insights and discoveries by leveraging the large volume of data available within the CDM. This can support collaborative research projects, benchmarking, and quality improvement initiatives. In addition, it can provide a more comprehensive view of patient health and facilitate research.

Finally, a good combination of HL7 FHIR with OMOP CDM can offer many benefits to the system like data exchange, facilitates research collaborations, and enables evidence generation using standardized and harmonized healthcare data. Using these standards together promotes data standardization, enabling interoperability and data exchange across different healthcare systems and research institutions. Researchers can leverage the rich clinical and demographic data available in FHIR resources, map them to OMOP CDM entities, and conduct standardized analyses and research studies using the OMOP CDM's data structure. This combination enables large-scale observational research and comparative effectiveness studies across multiple healthcare organizations. By aligning the OMOP CDM with FHIR, healthcare organizations can integrate their EHR systems with the OMOP CDM and leverage the FHIR-based APIs (Application Programming Interfaces) to access and exchange standardized clinical data in the OMOP CDM format.

This interoperability facilitates data sharing for research collaborations, quality improvement initiatives, and population health management.

6.2.7. Regarding standardize DB with archetypes to present data in the UI

Many times, DBBs are designed with only data exchange or processing and storage in mind, but rarely are they designed to be grouped according to templates or use cases that will need to be displayed in the user interface (UI).

EHR information systems tend to be quite similar around the world in terms of UI, functionality, and services they offer to healthcare professionals. It is for this reason that Open Electronic Health Record (OpenEHR)[89] seems to be a very suitable standard to design the DBB considering the functional needs that will be at the level of the user interface.

OpenEHR is an open-source initiative, driven by a global community of clinicians, researchers, and developers. It provides flexibility in data modelling, allowing organizations to adapt the standard to their specific healthcare domains and local requirements. It enables customization and extension of the information model to accommodate various clinical specialties, regional variations, and specific organizational needs.

[89] OpenEHR Overview. (n.d.). from https://openehr.org/OpenEHR vs HL7 FHIR (n.d.). https://vico.orghttps://vicoacademy.com/innova@vico.org

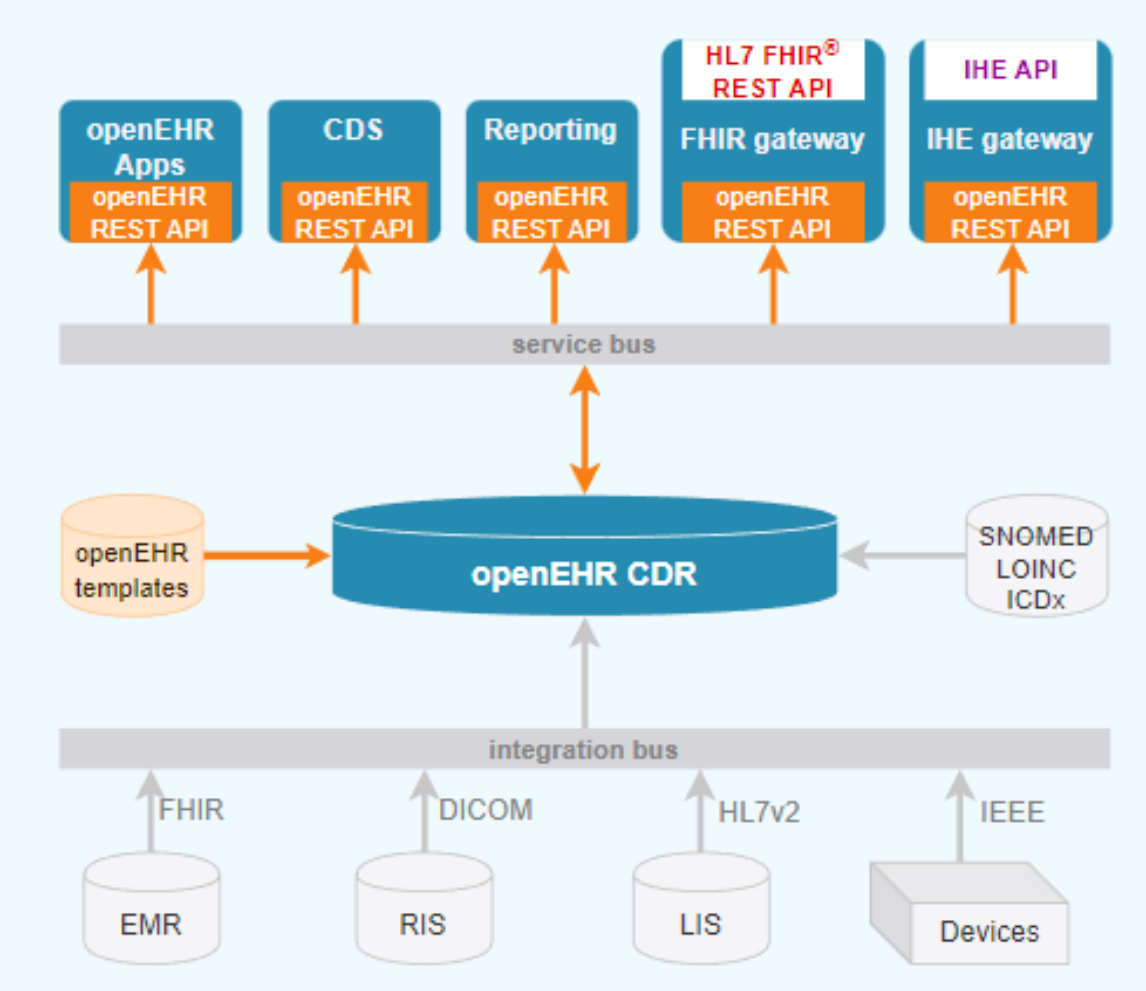


Figure 48 OpenEHR structure.

Also, it offers a more comprehensive approach to clinical modelling, enabling the representation of complex clinical concepts and relationships between data elements. It allows for the creation of archetypes, which are formalized specifications of clinical concepts, facilitating the precise representation of healthcare data. This granularity and clinical focus make this standard suitable for scenarios that require detailed clinical data modelling.

In summary, OpenEHR emphasizes semantic interoperability, focusing on representing healthcare information with a standardized information model, facilitates the exchange and interpretation of data across different systems and domains, provides a more comprehensive clinical modelling approach, supports the concept of a longitudinal health record, capturing and managing a patient's health information over time, utilizes archetypes and accommodate various healthcare domains and local requirements.

In this case mapping the data from HL7 FHIR to OpenEHR, or HL7 v2 o v3 to OpenEHR could be a good approach to display the data in any EHR information system.

In conclusion, when we think about architecture, information models, interoperability between different systems and standards, it is very important to choose the right standards for the definition of use cases and healthcare process, data exchange without losing the semantic meaning with terminologies, the exchange of clinical data and medical images, the exchange of clinical documents and the resolution of events, data indexing and clustering that give fast real-time answers and facilitate data analysis and research, and the correct representation of these in a standardized way in the UI.

It seems that the combination of standards with other standards is a highly recommended option for this sector. For example, an information system that combines the following standards in the following way, could be a very powerful interoperable system in the health sector: that the DBB is with OpenEHR to display the data in the UI; HL7 FHIR for the exchange of data, messaging, documents, auditing, etc.; DICOM and PACS for the exchange of medical images, DICOM SR for the exchange of image annotations, etc.; OMOP for data indexing and for statistical systems, analysis, prediction, etc.; all under the framework of IHE; making use of semantic standards with controlled vocabularies and terminologies such as LOINC, SNOMED CT, ICD, etc.

In any case, as it has already been commented several times in this document, technology and standards are constantly evolving and therefore, before making any decision, it is important to make a state of the art analysis and a study of this kind of discussions ongoing, to verify whether the standards have changed or evolved.

7. References

AJCC staging system - NCI Dictionary of Cancer Terms - NCI. (n.d.). from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/ajcc-staging-system
Algolia Overview. (n.d.). from https://www.algolia.com/, https://www.algolia.com/doc/
ATC Overview. (n.d.). from https://www.who.int/tools/atc-ddd-toolkit/atc-classification
ATNA IHE Overview. (n.d.). from https://wiki.ihe.net/index.php/Audit_Trail_and_Node_Authentication
CDM Overview. Weeks, J., & Pardee, R. Learning to Share Health Care Data: A Brief Timeline of Influential Common Data Models and Distributed Health Data Networks in U.S. Health Care Research. EGEMs. 2019 Mar; 7(1), 4. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6437693/
CDS Overview. Wasylewicz, A. T. M., & Scheepers-Hoeks, A. M. J. W. Clinical Decision Support Systems. Fundamentals of Clinical Data Science. 2018 Dec; 153–169. from https://www.ncbi.nlm.nih.gov/books/NBK543516/
CEF Overview. (n.d.). from https://commission.europa.eu/funding-tenders/find-funding/eu-funding-programmes/connecting-europe-facility_en
Common Guidelines for eHealth Harmonisation and Interoperability. (n.d.). from https://eufordigital.eu/wp-content/uploads/2021/03/Common-Guidelines-for-eHealth-Harmonisation-and-Interoperability.pdf
CPT Overview. Dotson, P. CPT® Codes: What Are They, Why Are They Necessary, and How Are They Developed? Advances in Wound Care. 2013 Dec; 2(10) from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3865623/
CVX Overview. (n.d.). from https://www2a.cdc.gov/vaccines/iis/iisstandards/vaccines.asp?rpt=cvx, https://www.nlm.nih.gov/research/umls/rxnorm/sourcereleasedocs/cvx.html
Demographic Information Model. (n.d.). from https://specifications.openehr.org/releases/RM/Release-1.1.0/demographic.html#_demographic_information_model
DGA Overview. (n.d.). from https://www.european-data-governance-act.com/
DICOM Overview. (n.d.). from https://www.dicomstandard.org/current
DICOMweb (n.d.). from https://www.dicomstandard.org/using/dicomweb
Difference beetween IHE XDS, HL7 FHIR, HL7 CDA. (n.d.). Retrieved July 26, 2023, from http://www.ringholm.com/column/combining_ihe_xds_mhd_and_fhir.htm, http://www.ringholm.com/column/bye_bye_IHE_XDS_and_CDA.htm
DIMSE Overview. (n.d.). from https://dicom.nema.org/medical/dicom/current/output/chtml/part07/sect_7.5.html

Documentation:vocabulary [Observational Health Data Sciences and Informatics]. (n.d.). from https://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary, https://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary:introduction
EHDS Overview. (n.d.). From https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
EHR Overview. Ambinder, E. P. (2005). Electronic Health Records. Journal of Oncology Practice. 2005 Jul; 1(2), 57. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2793588/
Elasticsearch Platform Overview(n.d.). from https://www.elastic.co/
EMDN Overview. (n.d.). from https://webgate.ec.europa.eu/dyna2/emdn/
ETL Overview. Cheng, K. Y., Pazmino, S., & Schreiweis, B. ETL Processes for Integrating Healthcare Data - Tools and Architecture Patterns. Studies in Health Technology and Informatics. 2022 Nov; 299, 151–156. from https://pubmed.ncbi.nlm.nih.gov/36325856/
EUDAMED Overview. (n.d.). from https://ec.europa.eu/tools/eudamed/#/screen/home
eXtensible Markup Language Introduction. (n.d.). from https://www.w3.org/XML/
FDA Overview. Technology, I. of M. (US) C. on H. C., & Goodman, C. Food and Drug Administration. from https://www.fda.gov/, https://www.ncbi.nlm.nih.gov/books/NBK218462/
FHIR collaboration platform - SIMPLIFIER.NET Overview. (n.d.). from https://simplifier.net/
FHIR Documents (n.d.). from https://build.fhir.org/documents.html
FHIR Messaging. from https://hl7.org/fhir/R4/messaging.html
FHIR RESTful API. (n.d.). from https://hl7.org/fhir/R4/http.html
FHIR Validator. (n.d.). from https://simplifier.net/validate?fhirVersion=R4&scope=hl7.fhir.r4.core@4.0.1
Forge Overview. from https://fire.ly/products/forge/
GDPR Overview. (n.d.). from https://edps.europa.eu/data-protection/our-work/subjects/health_en, https://pubmed.ncbi.nlm.nih.gov/32570555/
GitHub Overview (n.d.). from https://github.com/
HAPI FHIR. (n.d.). from http://hapi.fhir.org/, https://hapifhir.io/
HL7 CDA Overview. (n.d.). from http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
HL7 FHIR Documents (n.d.). 2023, from https://build.fhir.org/documents.html
HL7 FHIR Implementation Guide: Breast Cancer Data, Release 1 - US Realm (Draft for Comment 2). (n.d.). from http://hl7.org/fhir/us/breastcancer/2018Sep/
HL7 FHIR Overview. (n.d.). from https://www.hl7.org/fhir/index.html
HL7 FHIR Resourcelist - FHIR v4.0.1. (n.d.). from https://hl7.org/fhir/R4/resourcelist.html
HL7 Overview. (n.d.). from https://www.hl7.org/index.cfm
HL7 v2, HL7 v3 Overview. (n.d.). from https://www.hl7.org/implement/standards/product_section.cfm?section=13, https://www.hl7.org/implement/standards/product_section.cfm?section=14

HL7.FHIR.US.MCODE\Home - FHIR v4.0.1. (n.d.). from https://build.fhir.org/ig/HL7/fhir-mCODE-ig/
HL7FHIR_StateArt. (n.d.). https://vico.org/aFHIRbest/StateArt/HL7FHIR_StateArt.pdf
Human Phenotype Ontology. (n.d.). from https://hpo.jax.org
IBM FHIR Server. (n.d.). from https://www.ibm.com/downloads/cas/EWNW6QE3?mhsrc=ibmsearch_a&mhq=FHIR%20Server
ICD Overview. (n.d.). from https://www.who.int/standards/classifications/classification-of-diseases
IEC 62304:2006 - Medical device software — Software life cycle processes. (n.d.). from https://www.iso.org/standard/38421.html
IEEE Devices Overview. (n.d.). from https://standards.ieee.org/practices/healthcare-life-sciences/wamiii/
IHE Enterprise Document Sharing Overview. (n.d.). from https://wiki.ihe.net/index.php/Cross-Enterprise_Document_Sharing
IHE ITI TF Vol1. (n.d.). from https://profiles.ihe.net/ITI/TF/Volume1/ch-9.html, https://profiles.ihe.net/ITI/TF/Volume1/ch-10.html#10.4
IHE Overview. (n.d.). from https://www.ihe.net/
Implementation Guide Registry. (n.d.). from http://fhir.org/guides/registry/
INCISIVE - SIMPLIFIER.NET. (n.d.). from https://simplifier.net/INCISIVE
INCISIVE HL7 FHIR message examples (n.d.). from https://simplifier.net/incisive/~resources?category=Example&exampletype=Bundle&sortBy=RankScore_desc
IOD Overview. (n.d.). from https://dicom.nema.org/medical/dicom/current/output/chtml/part04/chapter_6.html
ISO 14971:2019 - Medical devices — Application of risk management to medical devices. (n.d.). from https://www.iso.org/standard/72704.html
ISO 27799:2016 Standard: Health informatics — Information security management in health using ISO/IEC 27002. (n.d.). from https://www.iso.org/standard/62777.html
ISO and IEC International Standards for policy makers. from https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100359.pdf
ISO Overview. (n.d.). from https://www.iso.org/iso-45001-occupational-health-and-safety.html
ISO/IEC 27001 Standard: Information Security Management Systems. (n.d.). from https://www.iso.org/standard/27001
IVDR Overview. (n.d.). from https://www.fda.gov/medical-devices/ivd-regulatory-assistance/overview-ivd-regulation
JavaScript Object Notation Introduction. (n.d.). from https://www.json.org/json-en.html
LOINC Overview. (n.d.). from https://loinc.org/

MDC Overview. (n.d.). from https://www.mdc-ce.de/health-care.html?L=0, https://www.mdc-ce.de/services.html?L=0
MDR Overview. (n.d.). from https://single-market-economy.ec.europa.eu/single-market/european-standards/harmonised-standards/medical-devices_en
MEDCIN Overview. (n.d.). from https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/MEDCIN/index.html
Medical Device Regulation (MDR) | TÜV SÜD. (n.d.). from https://www.tuvsud.com/en/industries/healthcare-and-medical-devices/medical-devices-and-ivd/medical-device-market-approval-and-certification/medical-device-regulation
MED-RT Overview. (n.d.). from https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/MED-RT/index.html
Nadeau, E. (2020). Nist privacy framework: a tool for improving privacy through enterprise risk management. https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework_V1.0.pdf
NANDA Overview. (n.d.). from https://nanda.org/
NDC Overview. (n.d.). from https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory
NDF-RT Overview. (n.d.). from https://evs.nci.nih.gov/ftp1/NDF-RT/NDF-RT%20Documentation.pdf
NHS Overview. (n.d.). from https://www.england.nhs.uk/nhs-standard-contract/previous-nhs-standard-contracts/22-23/
NIFTI Overview. (n.d.). from https://nifti.nimh.nih.gov/
OASIS Overview. (n.d.). from https://www.oasis-open.org/policies-guidelines/interoperability-guidelines/
OAuth Overview. (n.d.). from https://oauth.net/2/
OHDSI Overview. (n.d.). from https://www.ohdsi.org/
OMOP Overview. (n.d.). from https://www.ohdsi.org/data-standardization/, https://www.ohdsi.org/2019-tutorials-omop-common-data-model-and-standardized-vocabularies/, https://ohdsi.github.io/CommonDataModel/
OMOPonFHIR. (n.d.). from http://omoponfhir.org/#, https://github.com/omoponfhir/omoponfhir-site-n-docs/wiki
OpenEHR Overview. (n.d.). from https://openehr.org/OpenEHR vs HL7 FHIR (n.d.). https://vico.orghttps://vicoacademy.com/innova@vico.org
Orthanc - DICOM Server. (n.d.). from https://www.orthanc-server.com/
Orthanc REST API. (n.d.). from https://api.orthanc-server.com/
PACS Overview. Arora D, Mehta Y. Use of picture archiving and communication system for imaging of radiological films in cardiac surgical intensive care unit. J Anaesthesiol Clin Pharmacol. 2014 Jul;30(3):447-8. from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4152706/

Profiles - IHE International. (n.d.). from https://www.ihe.net/resources/profiles/
Radiology Technical Committee, I. (2023). IHE_RAD_TF_Rev21-0_Vol2_FT_2023-06-15. https://www.ihe.net/uploadedFiles/Documents/Radiology/IHE_RAD_TF_Vol2.pdf
ReEIF Overview. (n.d.). from https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b56dffdc&appId=PPGMS
RxNav Overview. (n.d.). from https://lhncbc.nlm.nih.gov/RxNav/applications/RxNavDoc.html?_gl=1*1gaq85j*_ga*MTE3NjM2Nzk0Ny4xNjg5NjgxNzg2*_ga_P1FPTH9PL4*MTY4OTc3NTU4Mi4yLjAuMTY4OTc3NTU4Mi4wLjAuMA..*_ga_7147EPK006*MTY4OTc3NTU4Mi4yLjAuMTY4OTc3NTU4Mi4wLjAuMA..
RxNorm. (n.d.). from https://www.nlm.nih.gov/research/umls/rxnorm/index.html
Segmentation IOD. (n.d.). from https://dicom.nema.org/dicom/2013/output/chtml/part03/sect_A.51.html
SNOMED CT - IHTSDO/snowstorm: Scalable SNOMED CT Terminology Server using Elasticsearch. (n.d.). from https://github.com/IHTSDO/snowstorm/tree/master
SNOMED CT Overview. (n.d.). from https://www.snomed.org/
SOP Overview. (n.d.). from https://www.dicomlibrary.com/dicom/sop/
SR Overview. (n.d.). from https://dicom.nema.org/dicom/2013/output/chtml/part20/sect_A.3.html
Technical Frameworks - IHE International. (n.d.). from https://www.ihe.net/resources/technical_frameworks/
UCUM Overview. (n.d.). from https://ucum.org/
UDI-DI Overview. (n.d.). from https://health.ec.europa.eu/medical-devices-topics-interest/unique-device-identifier-udi_es
UNE-EN 60601-1:2008 Equipos electromédicos. Part 1: Requisito... (n.d.). from https://www.en.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0041083
UNII Overview. (n.d.). from https://www.fda.gov/industry/structured-product-labeling-resources/uniis-preferred-substance-names-and-their-identified-synonyms
WHO Overview. (n.d.). from https://www.who.int/

ANNEX 1. INCISIVE SNOMED CT & LOINC encoded terms

This annex contains all SNOMED CT and LOINC terms used by INCISIVE.

1.1 SNOMED CT Terms Table

CODE	DESCRIPTION
18629005:363703001=373846009	|Administration of drug or medicament (procedure)|:|Has intent (attribute)|=|Adjuvant - intent (qualifier value)|
18629005:363703001=373847000	|Administration of drug or medicament (procedure)|:|Has intent (attribute)|=|Neo-adjuvant - intent (qualifier value)|
371480007:718497002=34402009	|Anatomic location of neoplasm (observable entity)|:|Inherent location (attribute)|=|Rectum structure (body structure)|
897713009:370130000=410672004	|Antineoplastic chemoimmunotherapy (regime/therapy)|:|Property (attribute)|=|Date property (qualifier value)|
897713009:370130000=410672004, 370134009=6493001	|Antineoplastic chemoimmunotherapy (regime/therapy)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=|Recent (qualifier value)|
897713009:370130000=410656007	|Antineoplastic chemoimmunotherapy (regime/therapy)|:|Property (attribute)|=|Type (property) (qualifier value)|
86273004:370130000=410672004	|Biopsy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
86273004:370130000=410656007	|Biopsy (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
86273004:370134009=7389001	|Biopsy (procedure)|:|Time aspect (attribute)|=|Time frame (qualifier value)|
367336001:370130000=410672004	|Chemotherapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
367336001:370130000=410672004, 370134009=6493001	|Chemotherapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=|Recent (qualifier value)|
367336001:370130000=410656007	|Chemotherapy (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
230056004:246514001=259032004	|Cigarette consumption (observable entity)|:|Units (attribute)|=|per day (qualifier value)|
703423002:370130000=410672004	|Combined chemotherapy and radiation therapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
703423002:370130000=410672004, 370134009=6493001	|Combined chemotherapy and radiation therapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=|Recent (qualifier value)|
703423002:370130000=410656007	|Combined chemotherapy and radiation therapy (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
371892002:704327008=76752008, 246514001=229029004	|Delivered radiation dose (observable entity)|:|Direct site (attribute)|=|Breast structure (body structure)|, |Units (attribute)|=|Gray (qualifier value)|

371892002:704327008=59441001, 246514001=229029004	|Delivered radiation dose (observable entity)|:|Direct site (attribute)|=|Structure of lymph node (body structure)|, |Units (attribute)|=|Gray (qualifier value)|
371892002:704327008=1162492000, 246514001=229029004	|Delivered radiation dose (observable entity)|:|Direct site (attribute)|=|Tumor bed (morphologic abnormality)|, |Units (attribute)|=|Gray (qualifier value)|
275937001:408732007=66839005	|Family history of cancer (situation)|:|Subject relationship context (attribute)|=|Father (person)|
275937001:408732007=72705000	|Family history of cancer (situation)|:|Subject relationship context (attribute)|=|Mother (person)|
308273005:255234002=387713003	|Follow-up status (finding)|:|After (attribute)|=|Surgical procedure (procedure)|
308273005:255234002=862730004	|Follow-up status (finding)|:|After (attribute)=|Biopsy (procedure)|
169413002:370130000=410672004	|Hormone therapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
169413002:370130000=410656007	|Hormone therapy (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
365853002:418775008=113091000	|Imaging finding (finding)|:|Finding method (attribute)|=|Magnetic resonance imaging (procedure)|
76334006:370130000=410672004	|Immunotherapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
76334006:370130000=410672004, 370134009=6493001	|Immunotherapy (procedure)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=Recent (qualifier value)|
76334006:370130000=410656007	|Immunotherapy (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
15220000:370134009=7389001	|Laboratory test (procedure)|:|Time aspect (attribute)|=|Time frame (qualifier value)|
364636000:719715003=76752008	|Lesion observable (observable entity)|:|Relative to part of (attribute)|=|Breast structure (body structure)|
113091000:370130000=410656007	|Magnetic resonance imaging (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
254837009:363714003=443941007	|Malignant neoplasm of breast (disorder)|:|Interprets (attribute)|=|Edition of American Joint Commission on Cancer, Cancer Staging Manual used for TNM staging (observable entity)|
363406005:363714003=443941007	|Malignant neoplasm of colon (disorder)|:|Interprets (attribute)|=|Edition of American Joint Commission on Cancer, Cancer Staging Manual used for TNM staging (observable entity)|
399068003:263502005=255318003	|Malignant tumor of prostate (disorder)|:|Clinical course (attribute)|=|Relapsing course (qualifier value)|
228862004:704321009=108290001	|Number of fractions (observable entity)|:|Characterizes (attribute)|=|Radiation oncology AND/OR radiotherapy (procedure)|
228862004:704327008=76752008, 704321009=108290001	|Number of fractions (observable entity)|:|Direct site (attribute)|=|Breast structure (body structure)|, |Characterizes (attribute)|=|Radiation oncology AND/OR radiotherapy (procedure)|

228862004:704327008=116249200, 704321009=108290001	|Number of fractions (observable entity)|:|Direct site (attribute)|=|Tumor bed (morphologic abnormality)|, |Characterizes (attribute)|=|Radiation oncology AND/OR radiotherapy (procedure)|
371497001:370130000=758637006	|pM category (observable entity)|:|Property (attribute)|=|Anatomic location (property) (qualifier value)|
63476009:370130000=410672004	|Prostate specific antigen measurement (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
385798007:370130000=758637006	|Radiation therapy care (regime/therapy)|:|Property (attribute)|=|Anatomic location (property) (qualifier value)|
385798007:370130000=410672004	|Radiation therapy care (regime/therapy)|:|Property (attribute)|=|Date property (qualifier value)|
385798007:370130000=410672004, 370134009=6493001	|Radiation therapy care (regime/therapy)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=|Recent (qualifier value)|
385798007:370130000=410656007	|Radiation therapy care (regime/therapy)|:|Property (attribute)|=|Type (property) (qualifier value)|
118247008:418775008=77477000	|Radiologic finding (finding)|:|Finding method (attribute)|=|Computed tomography (procedure)|
118247008:418775008=168537006	|Radiologic finding (finding)|:|Finding method (attribute)|=|Plain radiography (procedure)|
118247008:418775008=764685004	|Radiologic finding (finding)|:|Finding method (attribute)|=|Positron emission tomography of whole body using choline C-11 (procedure)|
118247008:418775008=450436003	|Radiologic finding (finding)|:|Finding method (attribute)|=|Positron emission tomography with computed tomography (procedure)|
118247008:418775008=228084000	|Radiologic finding (finding)|:|Finding method (attribute)|=|Whole body bone imaging (procedure)|
396152005:726633004=866310001	|Raised prostate specific antigen (finding)|:|Temporally related to (attribute)|=|Absence of signs (finding)|
396152005:726633004=118226009	|Raised prostate specific antigen (finding)|:|Temporally related to (attribute)|=|Temporal finding (finding)|
372280006:704325000=251695000	|Ratio of number of prostatic tissue cores positive for carcinoma to total number of cores obtained by needle biopsy (observable entity)|:|Relative to (attribute)|=|Structure of left lateral lobe of prostate (body structure)|
372280006:704325000=251597001	|Ratio of number of prostatic tissue cores positive for carcinoma to total number of cores obtained by needle biopsy (observable entity)|:|Relative to (attribute)|=|Structure of right lateral lobe of prostate (body structure)|
387713003:363704007=39607008	|Surgical procedure (procedure)|:|Procedure site (attribute)|=|Lung structure (body structure)|
387713003:370130000=410672004	|Surgical procedure (procedure)|:|Property (attribute)|=|Date property (qualifier value)|
387713003:370130000=410656007	|Surgical procedure (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
6021000124103:370130000=410672004	|Targeted medication therapy review (procedure)|:|Property (attribute)|=|Date property (qualifier value)|

6021000124103:370130000=410672004, 370134009=6493001	|Targeted medication therapy review (procedure)|:|Property (attribute)|=|Date property (qualifier value)|, |Time aspect (attribute)|=|Recent (qualifier value)|
6021000124103:370130000=410656007	|Targeted medication therapy review (procedure)|:|Property (attribute)|=|Type (property) (qualifier value)|
228487000:246514001=258707000	|Total time smoked (observable entity)|:|Units (attribute)|=|year (qualifier value)|
314705003:255234002=308273005	|Treatment plan given (finding)|:|After (attribute)|=|Follow-up status (finding)|
420415002	Acoustic feature of mass (observable entity)
423493009	Age at diagnosis (observable entity)
1221000175102	Age at smoking cessation (observable entity)
88111009	Altered bowel function (finding)
1222590007	American Joint Committee on Cancer pathological N category allowable value (qualifier value)
1222589003	American Joint Committee on Cancer pathological T category allowable value (qualifier value)
1229859000	American Joint Committee on Cancer pT3 (qualifier value)
1229861009	American Joint Committee on Cancer pT3b (qualifier value)
371480007	Anatomic location of neoplasm (observable entity)
707266006	Androgen deprivation therapy (procedure)
897713009	Antineoplastic chemoimmunotherapy (regime/therapy)
246090004	Associated finding (attribute)
86273004	Biopsy (procedure)
365855009	Biopsy finding (finding)
86273004:370134009=57615005	Biopsy:Time aspect=Definite time
445232009	Boost radiation therapy (procedure)
119184005	Breast part (body structure)
2483006	Cavity (morphologic abnormality)
125363000	Central zone necrosis (morphologic abnormality)
367336001	Chemotherapy (procedure)
404684003	Clinical finding (finding)
399387003	cM category (observable entity)
399534004	cN category (observable entity)
703423002	Combined chemotherapy and radiation therapy (procedure)
77477000	Computed tomography (procedure)
246191002	Consistency (attribute)
49727002	Cough (finding)
443987001	Cribriform neoplasm pattern (finding)
399504009	cT category (observable entity)
410671006	Date (attribute)
840417007	Date laboratory test due (observable entity)
439272007	Date of procedure (observable entity)
371892002	Delivered radiation dose (observable entity)
16310003	Diagnostic ultrasonography (procedure)

64572001	Disease (disorder)
11971000224104	Dissection of lymph node (procedure)
267036007	Dyspnea (finding)
49158009	Emphysema (morphologic abnormality)
364699009	Ethnic group (observable entity)
58347006	Excision of lymph node (procedure)
372306004	Extraprostatic extension of tumor present (finding)
414205003	Family history of prostate cancer (situation)
442301001	Finding of change compared to previous radiologic examination (finding)
263495000	Gender (observable entity)
372278000	Gleason score (observable entity)
66857006	Hemoptysis (finding)
396659000	Histologic grade of urothelial carcinoma by World Health Organization and International Society of Urological Pathology technique (observable entity)
250537006	Histopathology finding (finding)
252416005	Histopathology test (procedure)
417662000	History of clinical finding in subject (situation)
169413002	Hormone therapy (procedure)
76334006	Immunotherapy (procedure)
62112002	Injury of breast (disorder)
81060008	Intestinal obstruction (disorder)
81060008	Intestinal obstruction (disorder)
87522002	Iron deficiency anemia (disorder)
87522002	Iron deficiency anemia (disorder)
165331007	Laboratory procedure performed (situation)
15220000:370134009=57615005	Laboratory test:Time aspect=Definite time
272741003	Laterality (attribute)
263605001	Length dimension of neoplasm (observable entity)
364109001	Length of lymph node in excised specimen (observable entity)
364109001	Length of lymph node in excised specimen (observable entity)
246267002	Location (attribute)
17785005	Lung volume, function (observable entity)
443097008	Lymph node station of involved lymph nodes (observable entity)
443097008	Lymph node station of involved lymph nodes (observable entity)
277208005	M category (observable entity)
113091000	Magnetic resonance imaging (procedure)
113091000	Magnetic resonance imaging (procedure)
254837009	Malignant neoplasm of breast (disorder)
363406005	Malignant neoplasm of colon (disorder)
781382000	Malignant neoplasm of colon and/or rectum (disorder)
399068003	Malignant tumor of prostate (disorder)
363351006	Malignant tumor of rectum (disorder)
1149295006	Malignant tumor staging (tumor staging)
129792006	Mammographic architectural distortion of breast (finding)

129715009	Mammographic breast composition finding (finding)
71651007	Mammography (procedure)
397137005	Mammography assessment finding (finding)
129771006	Mammography reference location (finding)
243121000	Medical therapy (procedure)
43440003	Melanocyte stimulating hormone releasing factor (substance)
161712005	Menopause, function (observable entity)
395538009	Microscopic specimen observation (finding)
277206009	N category (observable entity)
786838002	Nodule of lung (finding)
86616005	Non-infiltrating intraductal carcinoma (morphologic abnormality)
440425000	Number of births at term (observable entity)
228862004	Number of fractions (observable entity)
405930005	Number of tumor nodules (observable entity)
246097001	Patient status (attribute)
422549004	Patient-related Identification code (observable entity)
241000124106	Performance measure status (finding)
241000124106	Performance measure status (finding)
373873005	Pharmaceutical / biologic product (product)
60046008	Pleural effusion (disorder)
371497001	pM category (observable entity)
371497001	pM category (observable entity)
82918005	Positron emission tomography (procedure)
764685004	Positron emission tomography of whole body using choline C-11 (procedure)
1082901000112100	Primary malignant neoplasm of breast with axillary lymph node invasion (disorder)
93880001	Primary malignant neoplasm of lung (disorder)
93880001:363714003=443941007	Primary malignant neoplasm of lung (disorder):|Interprets (attribute)|=|Edition of American Joint Commission on Cancer, Cancer Staging Manual used for TNM staging (observable entity)|
399734001	Primary tumor size (observable entity)
118890000	Procedure on lymph node (procedure)
416342005	Procedure related observable (observable entity)
129125009	Procedure with explicit context (situation)
1163261007	Qualitative distribution of primary malignant neoplasm (observable entity)
108290001	Radiation oncology AND/OR radiotherapy (procedure)
385798007	Radiation therapy care (regime/therapy)
385798007	Radiation therapy care (regime/therapy)
363680008	Radiographic imaging procedure (procedure)
129737002	Radiographic lesion margin characteristics (finding)
129732008	Radiographic lesion shape finding (finding)
398007000	Regulatory notes (record artifact)
182985004	Response to treatment (situation)

182985004	Response to treatment (situation)
257915005	Sampling - action (qualifier value)
257915005	Sampling - action (qualifier value)
372294007	Seminal vesicle invasion by tumor present (finding)
77176002	Smoker (finding)
127790008	Staining method (procedure)
31094006	Structure of lobe of lung (body structure)
59441001	Structure of lymph node (body structure)
387713003	Surgical procedure (procedure)
63671000122101	Symptoms and signs observed (record artifact)
78873005	T category (observable entity)
6021000124103	Targeted medication therapy review (procedure)
276239002	Therapy (regime/therapy)
272103003	Time patterns (qualifier value)
445528004	Treatment changed (situation)
182991002:408731000=410586007	Treatment given:Temporal context=Specified time
371502004	Tumor border configuration (observable entity)
409769002	Tumor calcification (finding)
371500007	Tumor configuration (observable entity)
406084004	Tumor nodule site (observable entity)
419835002	Tumor progression (finding)
385356007	Tumor stage finding (finding)
254292007	Tumor staging (tumor staging)
370380004	Ultrasound scan finding (finding)
369734008	Venous (large vessel) extramural invasion by tumor present (finding)
89362005	Weight loss (finding)
89362005	Weight loss (finding)
228084000	Whole body bone imaging (procedure)

1.2 LOINC Terms Table

CODE	DESCRIPTION
46108-7	18q chromosome deletion [Identifier] in Blood or Tissue by Molecular genetics method Nominal
76625-3	Alanine aminotransferase [Enzymatic activity/volume] in Blood
13986-5	Albumin/Protein.total in 24 hour Urine by Electrophoresis
100019-9	ALK gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal
1783-0	Alkaline phosphatase [Enzymatic activity/volume] in Blood
1920-8	Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma
58483-9	BRAF gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal

21639-0	BRCA1 gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal
59041-4	BRCA1+BRCA2 gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal
71426-1	C reactive protein [Mass/volume] in Blood by High sensitivity method
49765-1	Calcium [Mass/volume] in Blood
6875-9	Cancer Ag 15-3 [Units/volume] in Serum or Plasma
24108-3	Cancer Ag 19-9 [Units/volume] in Serum or Plasma
2039-6	Carcinoembryonic Ag [Mass/volume] in Serum or Plasma
2093-3	Cholesterol [Mass/volume] in Serum or Plasma
2085-9	Cholesterol in HDL [Mass/volume] in Serum or Plasma
2089-1	Cholesterol in LDL [Mass/volume] in Serum or Plasma
81747-8	Chromosome region 6q22 rearrangements in Tissue by FISH
2143-6	Cortisol [Mass/volume] in Serum or Plasma
38483-4	Creatinine [Mass/volume] in Blood
2161-8	Creatinine [Mass/volume] in Urine
96271-2	DNA mismatch repair protein Msh3 [Presence] in Cancer specimen by Immune stain
21666-3	EGFR gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal
42783-1	ERBB2 gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal
85337-4	Estrogen receptor Ag [Presence] in Breast cancer specimen by Immune stain
2324-2	Gamma glutamyl transferase [Enzymatic activity/volume] in Serum or Plasma
41652-9	Glucose [Mass/volume] in Venous blood
59585-0	GSTP1 gene+APC gene methylation [Presence] in Tissue by Molecular genetics method
20570-8	Hematocrit [Volume Fraction] of Blood
718-7	Hemoglobin [Mass/volume] in Blood
17856-6	Hemoglobin A1c/Hemoglobin.total in Blood by HPLC
85319-2	HER2 [Presence] in Breast cancer specimen by Immune stain
18474-7	HER2 Ag [Presence] in Tissue by Immune stain
20448-7	Insulin [Units/volume] in Serum or Plasma
42788-0	Insulin-like growth factor-I receptor Ag [Presence] in Tissue by Immune stain
21703-4	KRAS gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal
11053-6	Lactate dehydrogenase [Enzymatic activity/volume] in Red Blood Cells
95218-4	Leukocytes [#/volume] in Stem cell product
98260-3	Lymphovascular invasion extent in Cancer specimen Qualitative
74489-6	Melan-A and Ki67 Ag [Identifier] in Tissue by Immune stain
100026-4	MET gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal
81708-0	Microsatellite markers exhibiting instability/Microsatellite instability markers assessed in Cancer specimen
42637-9	Natriuretic peptide B [Mass/volume] in Blood
83052-1	PD-L1 by clone 22C3 [Presence] in Tissue by Immune stain
83062-0	PIK3CA gene [VCF] in Cancer specimen by Sequencing
63419-6	PIK3CA gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal

26515-7	Platelets [#/volume] in Blood
75940-7	Potassium [Mass/volume] in Blood
85339-0	Progesterone receptor Ag [Presence] in Breast cancer specimen by Immune stain
15326-2	Prostate Adenoma Volume by derived from height, width and length (US)
2857-1	Prostate specific Ag [Mass/volume] in Serum or Plasma
52457-9	Provider Information
21734-9	RET gene mutations tested for in Blood or Tissue by Molecular genetics method Nominal
2947-0	Sodium [Moles/volume] in Blood
53899-1	TGFBR1 gene+TGFBR2 gene targeted mutation analysis in Blood or Tissue by Molecular genetics method
3015-5	Thyrotropin [Units/volume] in Blood
21739-8	TP53 gene mutations found [Identifier] in Blood or Tissue by Molecular genetics method Nominal
3043-7	Triglyceride [Mass/volume] in Blood
98981-4	Urate [Mass/volume] in Blood
20977-5	Urea [Mass/volume] in Blood

ANNEX 2. INCISIVE HL7 FHIR Message Example for each cancer

INCISIVE has created four HL7 FHIR message examples using bundle resources that provided by HL7 FHIR. These four FHIR message examples can be seen in the INCISIVE Simplifier[90].

- HL7FHIRMessageExampleBreast.xml
- HL7FHIRMessageExampleColorectal.xml
- HL7FHIRMessageExampleLung.xml
- HL7FHIRMessageExampleProstate.xml

In the FHIR message exemple, all term values are for reference only, it is just a example, and each term value needs to be filled in according to the real use cases.

[90] INCISIVE HL7 FHIR message examples (n.d.). from https://simplifier.net/incisive/~resources?category=Example&exampletype=Bundle&sortBy=RankScore_desc