EHTEL
Implementers Advancing Digital Health and Care Transformation

European-level requirements and constraints for FAIR data structures to enable communication between EHDS compliant EHR systems and algorithm-based tools

Final report

Version 2.4 of 14/01/2026

This report has been prepared in the context of the Xt-EHR Joint Action

Xt-EHR @dataspace

Co-funded by the European Union

My health @ EU
eHealth Digital Service Infrastructure
A service provided by the European Union

Table of Contents

Table of abbreviations

Abbreviation	**Full term**	**Definition of the term or the context in which it is used in the report**
AI	Artificial Intelligence	Systems using learning or adaptive algorithms to support prediction, decision-making or automation
AI Act	EU Artificial Intelligence Act	EU regulation governing high-risk and general-purpose AI systems
AI-CDSS	AI-supported Clinical Decision Support System	AI-enabled systems supporting clinical decision-making
AI-DSS	AI-supported Decision Support System	A broader category (than AI-CDSS) including clinical and non-clinical decision support
API	Application Programming Interface	Interface enabling software systems to exchange data
ATNA	Audit Trail and Node Authentication	IHE profile for security auditing and logging
BP	Blood Pressure	Clinical measurement used in remote monitoring
CDS/CDSS	Clinical Decision Support (Systems)	Systems providing knowledge and patient-specific information
CE	Conformité Européenne	Certification marking for medical devices in the EU
CHF	Congestive Heart Failure	Clinical use case described in detail
CI-SIS	Cadre d'Interopérabilité des Systèmes d'Information de Santé	French national interoperability framework

CIO	Chief Information Officer	Executive responsible for IT strategy and infrastructure
CMIO	Chief Medical Information Officer	Executive bridging clinical practice and digital systems
CPME	Standing Committee of European Doctors	A European medical professional organisation
CPOE	Computerised Physician Order Entry	Electronic ordering of medications and treatments
CSV	Comma-Separated Values	Machine-readable data format
CT	Computed Tomography	A medical imaging technique
DICOM	Digital Imaging and Communications in Medicine	Standard for medical imaging data
DICOM-WS	DICOM Whole Slide Imaging	A technology to digitally scan and archive whole slides in high resolution
DPIA	Data Protection Impact Assessment	GDPR-mandated privacy risk assessment
DSS	Decision Support System	Systems supporting decision-making processes
EDQM	Standard terms defined by the European Directorate for the Quality of Medicines & HealthCare	Terms and definitions to describe pharmaceutical dose forms
EEHRxF	European Electronic Health Record Exchange Format	Interoperability framework under the EHDS
EHDS	European Health Data Space	EU framework for primary and secondary use of health data
EHR	Electronic Health Record	A system used to store and manage electronic health data
EMR	Electronic Medical Record	Often used synonymously with EHR in a usability context
ETL	Extract-Transform-Load	Data integration and processing pipeline
EU	European Union	Regulatory and policy context
FAIR	Findable, Accessible, Interoperable, Reusable	Data principles underpinning the report
FHIR (FHIR®)	Fast Healthcare Interoperability Resources	HL7 standard for health data exchange
GDPR	General Data Protection Regulation	EU data protection regulation
GP	General Practitioner	Primary care clinician in diabetes use case
HAS	Haute Autorité de Santé (France)	HAS is tasked with advancing quality in health and social care
HIS	Hospital Information System	Element of health informatics that primarily addresses the administrative and operational needs of hospitals
HbA1c	Glycated Haemoglobin	Clinical biomarker for diabetes management
HFrEF	Heart Failure with Reduced Ejection Fraction	CHF subtype used in the use case
HL7	Health Level Seven	International healthcare interoperability standards body
HIMSS	Healthcare Information and Management Systems Society	Organisation defining EHR usability
HTA	Health Technology Assessment	EU-level evaluation of clinical and economic value

HR	Heart Rate	Mentioned in CHF use case
ICD	International Classification of Diseases	WHO clinical coding system
ICPC	International Classification of Primary Care	Primary care coding system
IEEE	Institute of Electrical and Electronics Engineers	Standards body (e.g. IEEE 11073 for devices)
IHE	Integrating the Healthcare Enterprise	Interoperability profiles organisation
IoT	Internet of Things	Connected devices generating health data
ISO	International Organization for Standardization	Standards body
ISO IDMP	Identification of Medicinal Products	Standard for medicinal product data
IT	Information Technology	Technical systems and infrastructure
ITL	IT / EHR Integration Lead	Role responsible for system integration
JSON	JavaScript Object Notation	Data exchange format
LLM	Large Language Model	Advanced AI model used for NLP
LOINC	Logical Observation Identifiers Names and Codes	Standard for lab and clinical observations
MDR	Medical Device Regulation	EU regulation for medical devices
NICE	National Institute for Health and Care Excellence (UK)	Produces useful and usable guidance for the NHS and wider health and care system
NLP	Natural Language Processing	AI techniques for unstructured text
NIS2	Network and Information Security Directive (v2)	EU cybersecurity regulation
NIST	National Institute of Standards and Technology	US standards body cited in footnote 3
OMOP	Observational Medical Outcomes Partnership	Common data model
OAuth	Open Authorization	Secure authorisation protocol
PACS	Picture Archiving and Communication System	Medical imaging system
PCCP	Predetermined Change Control Plan	A regulatory concept (which is not allowed under the MDR)
PHR	Personal Health Record	Patient-controlled health record
PREMs	Patient-Reported Experience Measures	Patient experience data
PROMs	Patient-Reported Outcome Measures	Patient outcome data
PROV-O	Provenance Ontology	W3C standard for data lineage
RBAC	Role-Based Access Control	Security access model
RDF	Resource Description Framework	Linked Data representation format
REST	Representational State Transfer	API architectural style
RPM	Remote Patient Monitoring	Continuous monitoring outside of a hospital
SMART	Substitutable Medical Applications & Reusable Technologies	FHIR-based app integration framework
SNOMED CT	Systematized Nomenclature of Medicine – Clinical Terms	Clinical terminology system

SME	Small and Medium-sized Enterprise	AI health solution provider
SPARQL	SPARQL Protocol and RDF Query Language	Query language for RDF
UCUM	Unified Code for Units of Measure	Used in medication terminology
UI	User Interface	Means by which the user and a computer system interact
USA	United States of America	Referenced for comparative examples
URI	Uniform Resource Identifier	Linked Data identifier
UX	User Experience	Overall experience of a person using a digital product
W3C	World Wide Web Consortium	Web standards organisation
WP	Work Package	A group of tasks grouped together as a package. A term that is often used in EU projects
XML	Extensible Markup Language	Structured data format

Table of tables

Table of figures

1 Introduction: background and objectives

This section introduces the purpose, objective, and background of this report.

The first five sections of this report include descriptions of requirements for data fluidity (section 2) and for FAIR data (section 3), including a recommendation on organisational measures needed to implement FAIR data (sub-section 3.5); an outline of six overall tools to facilitate the FAIRification of data (section 4); and an overview of examples of current implementation of artificial intelligence (AI)-based algorithms in electronic health record (EHR) systems (section 5).

The following six sections (sections 6 to 11) focus more specifically on the FAIR requirements. Section 6 provided an overview of the main barriers to data fluidity, section 7 describes the high level functional and non-functional requirements, and section 8 illustrates those requirements through two use cases. Section 9 proposes a typology of the actors involved in these processes and suggests to use a framework to support them. Section 10 synthesises the priority development areas, while section 11 reflects on the use of synthetic data.

Two key areas of the report are its recommendations and conclusions. Section 12 elaborates 10 key recommendations which have resulted from the work undertaken for this report, before some conclusions are proposed in section 13.

Five annexes complement the report. Annex 1, called a typology, lists the 10 most used algorithm-based tools. Annex 2 offers a sample of 10 types of tools which are already available on the market. Annexes 3 and 4 provide detailed summaries of the two workshops associated with this report, organised in October and December 2025. Annex 5 provides links to the video recordings of the two workshops.

1.1 Purpose

EHTEL was mandated by Xt-EHR Joint Action to produce a report inspired by the remit of Xt-EHR WP5 deliverables.

The terms of reference for this report are defined in the following way (emboldening added): "the aim […] is to **define European-level expectations for data structures,** building on the results of previous EU-funded projects, **to enable communication between EHR systems and algorithm-based tools in line with the FAIR (Findable, Accessible, Interoperable, Reusable) data principles**.

Expected results: **European-level requirements and constraints for FAIR data structures to enable communication between EHDS compliant EHR systems and algorithm-based tools**."

Two workshops were foreseen to engage large audiences which would reflect on this topic. Before the workshops, a briefing paper was provided to all the people who had registered to

them. For the sake of consistency, elements of that paper have been incorporated in this report.

The briefing paper had as its main objective to prepare attendees to contribute actively to the two workshops by providing them with:

- An introduction to the concept of data fluidity and related requirements,
- An introduction to the possible organisational measures needed to produce FAIR data,
- An introduction to the tools which may help with structuring data in EHR systems and making them compliant with FAIR principles,
- A description of the possible contribution of validated authenticated sources for the FAIRification of data and the use of algorithm tools,
- A typology of the objectives of the algorithm-based tools used by EHR systems (Annex 1),
- A sample of tools currently deployed (Annex 2).

In particular, the second workshop provided feedback which has enabled the enrichment of the original briefing paper, and the expansion of this report to include a set of recommendations and conclusions.

1.2 The EHDS Regulation

The EHDS Regulation[1] provides for the first time a comprehensive interoperability framework for the healthcare domain which will ensure both a better continuity of care at national and European Union (EU) levels and a better access to and use of the data produced. It has also to be seen as an important enabler for the FAIRification of data and, hence, for the use of algorithm-based tools which will consume that data.

The interoperability component of the regulation should ensure a capacity to produce data compliant with the European Electronic Health Record Exchange Framework (EEHRxF) for the most frequently used health information domains.

Article 15.1 of the EHDS Regulation stipulates that “The Commission shall, by means of implementing acts, lay down the technical specifications to the priority categories of personal electronic health data referred to in Article 14(1), setting out the European electronic health record exchange format.”

[1] https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en

The format "shall include the following elements:

a) harmonised datasets containing electronic health data and defining structures, such as data fields and data groups for the representation of clinical content and other parts of the electronic health data;

b) coding systems and values to be used in datasets containing electronic health data;

c) technical interoperability specifications for the exchange of electronic health data, including its content representation, standards and profiles."

1.3 Definitions

This report refers to algorithm-based tools. Given the increasingly implicit association between algorithms and artificial intelligence (AI), it has been important to clarify some foundational definitions.

Therefore, in this report, we refer not only to algorithm-based tools but also use the terms AI (supported)-Decision Support Systems (AI-DSS) and AI (supported)-Clinical Decision Support System (AI-CDSS). For the sake of consistency, we differentiate conceptually between the notions of "algorithm" and "AI", which are in reality closely connected. Algorithms are specific sets of rules or instructions to solve a problem (e.g., sorting a list, finding a route) while AI is a broad field aiming to create systems that can learn, reason, and act intelligently, often by combining multiple algorithms, especially those that adapt data (such as machine learning algorithms). Algorithms are the core components that AI systems execute. AI uses sophisticated algorithms that can learn and can improve themselves (unlike simple, deterministic algorithms). An algorithm (or a set of algorithms) is considered to be AI when it can learn from user data and make complex predictions. (Not all algorithms are, however, supported by AI, as they can also be supported by fixed rules.)

AI-DSS refer to all processes – whether they are administrative, logistical, statistical or clinical – which are enhanced by algorithms that can learn and improve from data. When the term "clinical" is added to this definition, to become AI-CDSS, it refers specifically to the processes which support care for a specific patient, whether preventive, predictive or curative.

This report considers sophisticated algorithms supported by AI since there is now a clear trend towards their integration in EHR systems. However, most of the recommendations proposed (in Section 12 of this report) can be applied to both AI and non-AI supported algorithms.

When writing about Electronic Health Record (EHR) systems in this report, we refer to article 2 of the EHDS Regulation which define them as "any appliance or software intended by the manufacturer to be used for storing, intermediating, importing, exporting, converting, editing or viewing electronic health records". These systems must thus be understood in a broad

sense, to include systems used by healthcare organisations such as hospitals, individual healthcare professionals, and patients themselves.

2 Data fluidity: an introduction on requirements

Data fluidity (or data liquidity) is crucial for ensuring that algorithms are built on high-quality, diverse, and relevant datasets that are continuously updated to reflect real-world changes.

The concept of data fluidity can be defined as "the seamless movement and integration of data across organizational systems, ensuring timely access and informed decision-making[2]".

The six main requirements to achieve effective data fluidity can be summarised as follows:

1. **Standardisation of Data Formats and Protocols**: Implementing standardised data formats and communication protocols is crucial for interoperability among disparate health information systems. This standardisation facilitates consistent data exchange and integration across platforms.
2. **System Interoperability**: Developing technical architectures that support plug-and-play integration of devices and systems enables real-time, two-way data exchange. Such architectures enhance operational efficiency and care quality by allowing seamless communication between various healthcare technologies.
3. **Data Quality Assurance**: Ensuring data completeness, validity, consistency, uniqueness, timeliness, and accuracy is fundamental for reliable health information systems. High-quality data supports effective clinical decision-making and patient care.
4. **Legal and Regulatory Compliance**: Adhering to evolving legal frameworks is vital for maintaining data security and patient privacy. Healthcare organisations must stay updated about regulatory changes to ensure compliance and protect sensitive health information.
5. **User-Centred Design**: Understanding the needs and requirements of various stakeholders – including patients, clinicians, and healthcare providers – is essential for developing systems that are both effective and user-friendly. Engaging users in the design process helps tailor systems to meet real-world demands. Here, usability is key.
6. **Data Harmonisation**: Aligning and integrating data from diverse sources through harmonisation techniques enhances the comparability and usability of health

[2] This quote has been extracted from the text of a Dun & Bradstreet India author, Mukesh Kumar Jain: https://www.dnb.co.in/blog/why-fluidity-is-the-goal-to-mastering-your-data

information. This process supports comprehensive analysis and informed decision-making.

Four out of these six data fluidity requirements are covered by the EHDS or other European Regulations. The other two, which are not included in such regulations, are described briefly below.

2.1 Data quality assurance: a requirement for data fluidity

Data is the fuel, and AI is the engine. Data quality is the key to AI because the accuracy, reliability, and relevance of the outputs generated by these systems are directly dependent on the quality of the data that they are trained on; essentially, "garbage in, garbage out". Poor data leads to inaccurate and unreliable AI results, while high-quality data ensures that AI models learn accurate patterns and produces meaningful insights."[3]

In terms of **data quality assurance**, guidance is expected to emerge from the EU-funded QUANTUM project[4].

In February 2025, QUANTUM released two deliverables[5] covering aspects related to "specification of the data sets' quality and utility label". The label specification aims to cover the provisions of Article 78 of the EHDS Regulation.

Four key concepts are highlighted in those deliverables:

1. **Data quality** refers to features that describe the content of a dataset: typically, these are completeness, uniqueness, accuracy, validity, timeliness, and consistency of the data.
2. **Fit-for-use** is equivalent to a dataset's potential utility: typically, information that describes the dataset in a standard manner, such as source, provenance or lineage, content, or distribution.
3. **Fit-for-purpose** refers to the utility of an accessed dataset according to the objective of the health data request. The information necessary to evaluate fit-for-purpose depends on a user's experience. This information needs to be added to the description of a dataset as fit-for-purpose meta-data.
4. **Maturity** is a feature at the data holder level. It refers to the level of automation of data and data quality management procedures.

[3] https://www.nist.gov/itl/ssd/systems-interoperability-group/health-it-testing-infrastructure/health-it-testing-1
[4] https://quantumproject.eu/
[5] Specification for the assessment of data holders maturity and Specification of the data sets' quality and utility label

Although the part of the EHDS Regulation covering Article 78 refers to the secondary use of data, many of the concepts and principles described in the QUANTUM project guidance are also relevant to quality assurance for the primary use of data.

As a result, two concepts – **fit-for-use** and **maturity** – might be useful in order to set up requirements for EHR systems:

- The concept of fit-for-use might be applied to the capacity of an EHR system to interact with algorithm-based tools.
- The concept of maturity would be instrumental in 'qualifying' how a specific healthcare organisation has implemented a data quality management process.

2.2 User-centred design: another key requirement for data fluidity

Although **co-creation** is often mentioned as a key requirement by major initiatives and projects, there have been few initiatives or projects which have focused specifically on successfully implemented **user-centred design** as a method which facilitates the production of structured and quality data needed to interact with algorithm-based tools.

In describing the importance of user-centred design, it is evidently important to refer to **usability**, which forms an essential element of this type of design.

For example, the Healthcare Information and Management Systems Society (HIMSS) defines the usability of electronic medical record (EMR) as "the effectiveness, efficiency, and satisfaction with which specific users can achieve a specific set of tasks in a particular environment"[6]. The authors of this report consider that this EMR definition can also apply to EHR systems.

A quick analysis of the available literature shows that the general level of satisfaction of EHR users is relatively low (around 50%[7]). EHRs often need to be customised to support workflows which are specific to each organisation, and this often leads to important usability challenges.

In its policy paper "Implementing a 'user-friendly' European Health Data Space" of March 2025[8], the Standing Committee of European Doctors (CPME) reports that "Most EHR systems were created with billing and statistical objectives in mind, neglecting users. Studies show that EHR systems are not performing well in comparison with other product technologies." (p2) CPME claims therefore that "Usability measures must be adopted to evaluate effectiveness, efficiency and healthcare professional satisfaction when using an EHR system." (p1)

[6] https://www.himss.org/sites/hde/files/FileDownloads/HIMSS_DefiningandTestingEMRUsability.pdf
[7] See, for example https://pubmed.ncbi.nlm.nih.gov/31735343/ and https://pubmed.ncbi.nlm.nih.gov/33871018/ which focus on USA-based physicians and nurses.
[8] https://www.cpme.eu/news/implementing-a-user-friendly-and-intuitive-electronic-health-record-is-the-only-way-forward

Usability is often not part of certification processes and, even when it is included under certification, it is usually not applied to the specific onsite implementation (due, among other items, to a lack of appropriate resources and methods on the site).

Benefits for usability may result from the potential for interaction between EHR systems and algorithm-based tools to bring about important changes, so that more information and data become directly accessible to users. At the same time, a number of new usability challenges may also arise which need to be closely monitored and continuously adapted to.

3 FAIR data: an introduction on requirements

The concept of FAIR data is mentioned, although only twice, in the Preamble to the EHDS Regulation (notes 3 and 85) and is cited in one article only – Article 92.8 – where it relates to the working of the EHDS Board and the Board's cooperation with other relevant bodies "with a view to reaching advanced solutions towards findable, accessible, interoperable and reusable (FAIR) data usage in research and innovation". However, FAIR data structures can be considered as a precondition for making use of AI and other algorithm-based tools[9].

This section of the report offers an initial definition of FAIR data, important underpinning technical standards, means for achieving FAIR data, barriers to accomplishing FAIR data, and a recommendation regarding organisational measures needed.

3.1 What is meant by FAIR data

To enable communication between EHR systems and algorithm-based tools, the data structures used must also align with the **FAIR (Findable, Accessible, Interoperable, and Reusable)** data principles.

Here is a breakdown of the minimum expectations related to each principle:

1. Findable

- **Metadata Standards**: Use standardised and rich metadata (e.g., HL7[10], FHIR[11], DICOM[12]) to describe datasets and their relationships. This approach allows systems to locate and identify relevant data efficiently.
- **Unique Identifiers**: Ensure that every dataset, record, and version has a unique and persistent identifier (e.g., DOIs[13], ORCIDs[14]) for tracking purposes.

[9] It is worth mentioning that some AI-related tools could greatly facilitate the FAIRification of data and/or support a data quality assurance process.
[10] https://www.hl7.eu
[11] https://fhir.org
[12] https://www.dicomstandard.org
[13] https://www.doi.org/the-identifier/what-is-a-doi/
[14] https://orcid.org

- **Clear Indexing**: Data should be indexed with clearly defined attributes like date of collection, patient demographic details, and clinical context to facilitate searchability.

2. Accessible

- **Data Storage & APIs**: Data should be stored in a structured format (e.g., relational databases, NoSQL systems[15], cloud repositories) that allows for scalable access. RESTful APIs[16], or other open protocols like FHIR or OAuth[17], should be available to facilitate data access from external systems.
- **Standardised Data Formats**: Use widely accepted standards for data representation (e.g., JSON[18], XML[19], RDF[20]) so that data is readable across different systems and platforms.
- **Authentication & Authorisation**: Data access should be controlled, ensuring secure data exchange between EHRs and algorithmic tools, typically involving role-based access controls (RBAC[21]) and strong encryption.

3. Interoperable

- **Standardised Data Models**: Adopt standard data models, such as FHIR, HL7 for clinical data exchange, or OMOP[22] for observational health data. This approach allows seamless integration between systems by ensuring common terminology, units, and formats.
- **Data Harmonisation**: Map data to standard ontologies like SNOMED CT[23], ICD[24], or LOINC[25] to ensure that clinical terminology is consistent across different platforms and applications.
- **Semantic Interoperability**: Systems should interpret the meaning of data consistently, not just exchange raw data. A structured format, like FHIR, can define resource types for clinical data such as observations, medications, and procedures.

---

15 https://en.wikipedia.org/wiki/NoSQL
16 https://aws.amazon.com/what-is/restful-api/
17 https://en.wikipedia.org/wiki/OAuth
18 https://www.json.org/json-en.html
19 https://en.wikipedia.org/wiki/XML
20 https://www.w3.org/RDF/
21 https://en.wikipedia.org/wiki/Role-based_access_control
22 https://www.ohdsi.org/data-standardization/
23 https://www.snomed.org/what-is-snomed-ct
24 https://icd.who.int/en/
25 https://loinc.org

4. Reusable

- **Data Provenance**: Track data lineage and provenance so that the data's origin, transformations, and quality can be assessed. This can be done using open standards like PROV-O[26] (W3C Provenance Ontology).
- **Machine-readable Formats**: Use formats that are suitable for reuse, including those that are easily machine-readable by algorithmic tools (e.g., CSV[27], JSON, RDF). The data should include relevant clinical and technical details to facilitate analysis by automated algorithms.
- **Version Control**: Maintain versioning of data and algorithms, ensuring reproducibility and adaptability for future use cases. This can involve tracking data versions, updates to algorithms, and data schema changes.

3.2 Key technical standards to consider

At least seven key technical standards need to be considered when working on FAIR data. They are listed below[28].

- **HL7**: A framework for data exchange in healthcare, covering messaging and document exchange and more specifically **FHIR® (Fast Healthcare Interoperability Resources)**: It is a widely adopted standard for EHR and health data exchange.
- **DICOM**: Standard for medical imaging data.
- **OpenEHR**: Standard for electronic health record architectures.
- **OMOP**: Common data model for observational data analysis.
- **C-DISC**: A complete suite of standards, supporting clinical and non-clinical research processes from end to end.
- **Linked Data**: Use of RDF[29], SPARQL[30], and URIs[31] for representing clinical and algorithmic data to enable linking and data integration.
- **ISO IDMP**: A suite of standards that specify the use of standardised definitions for the identification and description of medicinal products for human use.

---

[26] https://www.w3.org/TR/prov-o/
[27] https://en.wikipedia.org/wiki/Comma-separated_values
[28] Observations on these standards (e.g., OpenEHR) are often located elsewhere in the report. As an example, OpenEHR is described in sub-sections 4.4 and 8.2, and Annex 2. SNOMED and SNOMED CT are referred to in sections 4, 5, 8, 12, and in both the workshops organised (Annexes 3 and 4).
[29] https://www.w3.org/RDF/
[30] https://www.w3.org/TR/sparql11-query/
[31] https://www.w3.org/Addressing/URL/uri-spec.html

By ensuring that data structures are aligned with these standards, communication between EHR systems and algorithm-based tools can be made more effective, efficient, and in line with the FAIR principles. This alignment enables better integration and should foster the potential for enhanced clinical decision-making, data reuse, and machine learning applications.

Aside from these technical data structure specifications, the use of controlled vocabularies (e.g. SNOMED CT, ICD, ICPC) is also a vital contribution to data quality assurance.

3.3 Data FAIRification: not only a technical challenge?

Achieving the implementation of FAIR data is more than simply a technical challenge. Three other aspects need to be considered:

- Ensuring data interoperability requires that each organisation develops a **FAIR data strategy** and allocates the necessary resources relevant both to human and technical efforts to reach this objective. Working with a performant EHR system is a necessary, but not sufficient, criterion.
- Data is produced locally for local primary use and is thus often based on implicit local conventions. To comply with FAIR principles, implicit **local conventions need to be made explicit** to enable data to be transformed into a fully specified, correct, and unambiguous data representation.
- At the organisational level, this transformation needs to be conducted in the most effective and efficient way, thus requiring the creation of an adapted, formalised, **semi-automated methodology**[32]. This methodology should always involve human experts – knowledge managers and data managers – who maintain automated processes and ensure the quality of the output and the availability of adapted and up-to-date tools. The methodology also needs to be agile in order to deep dive into every single data value and into natural language text expressed in multiple European and global languages.

3.4 Barriers to FAIR data

FAIR data structures are a precondition to make use of AI and other algorithm-based tools. There are, however, some barriers to the structures' use. Those barriers relevant to healthcare organisations are outlined below: the disconnect between coding teams and clinicians; the complexity and heterogeneity of systems; the role of external factors; the lack of traceability and explainability of systems; and the difficulties posed by one-shot implementations.

[32] See the recommendation below (3.5), listed among the tools to ensure the FAIRification of data.

Today, **hard coding of data** is still largely related to non-clinical objectives such as "pay for performance" policies. The work of hard coding is often done by dedicated technical teams whose members are disconnected from clinical practice. A majority of clinicians are reluctant to code data as they see it as time-consuming and as having an unclear value for them.

Considering the complexity of health systems, **heterogeneity** is common and details are often of major importance. Interoperability, however, requires that all implicit assumptions embedded in any data need to be understood, made explicit, and formalised.

Healthcare standards are essential for consistent communication, but they are not sufficient on their own because their **effectiveness depends on other external factors** like implementation, adaptability, and the human element of teamwork and communication. Standards enable clarity and prevent misunderstandings by providing a common language and framework, but failures arise when these standards are poorly designed, not updated in line with new technologies, or not consistently applied across different providers and contexts. Effective communication also requires human factors such as active team collaboration, situation awareness, and the ability to respond to the specific needs of each patient. Moving towards a **standardisation ecosystem**[33], where organisations are collaborating to develop, implement, learn from implementation and adapt standards, would be invaluable. Although considerable progress has been made on the use of standards, a lot remains to be done before the goal of a fully integrated standardisation ecosystem can be reached.

In the context of data interoperability, mappings, conversions, and translations need to be **traceable and explainable**. When considering the use of AI-driven algorithmic tools, **context-dependence** is also of paramount importance. The quality of an AI system can diminish or even be lost if it is used in a context which is different from the one in which the system was originally trained.

Data integration and interoperability cannot be solved by **one-shot implementations**. Instead, one needs to combine technological and methodological innovations: namely, the use of novel automated tools in a human-centric and end-to-end data integration methodology that clearly defines the actors and the steps at which the tools are activated.

3.5 A recommendation for organisational measures

A recommendation can therefore be made on how to approach the organisational aspects of setting up, curating, and maintaining FAIR data, especially with the help of a data manager.

[33] https://cerre.eu/wp-content/uploads/2023/05/230502_CERRE_Standardisation-report-.pdf

To overcome the organisational barriers to FAIR data:

- A robust methodology is needed to address the key requirements of both clinicians and researchers, in relationship to precision, responsibility, legal and regulatory compliance, explainability, and traceability.
- A common understanding (among clinicians, researchers, and IT personnel) of the need for FAIR data, and how to achieve it, is crucial. Only if the data are formally and unambiguously represented, down to the lowest level of individual data values, is it possible to apply automated transformations without the risk of corrupting health data.
- In order to maintain the quality of people's understanding of data, a **data manager** should put in operation a system which is based on three steps (see Figure 1[34]):
    - **set up** (configure) the system before the process; typically, the data manager would be working on a significant data sample (e.g. a thousand records) that is characteristic of the heterogeneity encountered in real life,
    - **curate** the results (or samples of the data) during the process,
    - **maintain** the system setup to follow the evolution of the requirements.

The proportion of automated and human interventions can be set up according to the level of criticality of mistakes (these proportions may differ in terms of the percentage of the importance given e.g., to care or research or accounting).

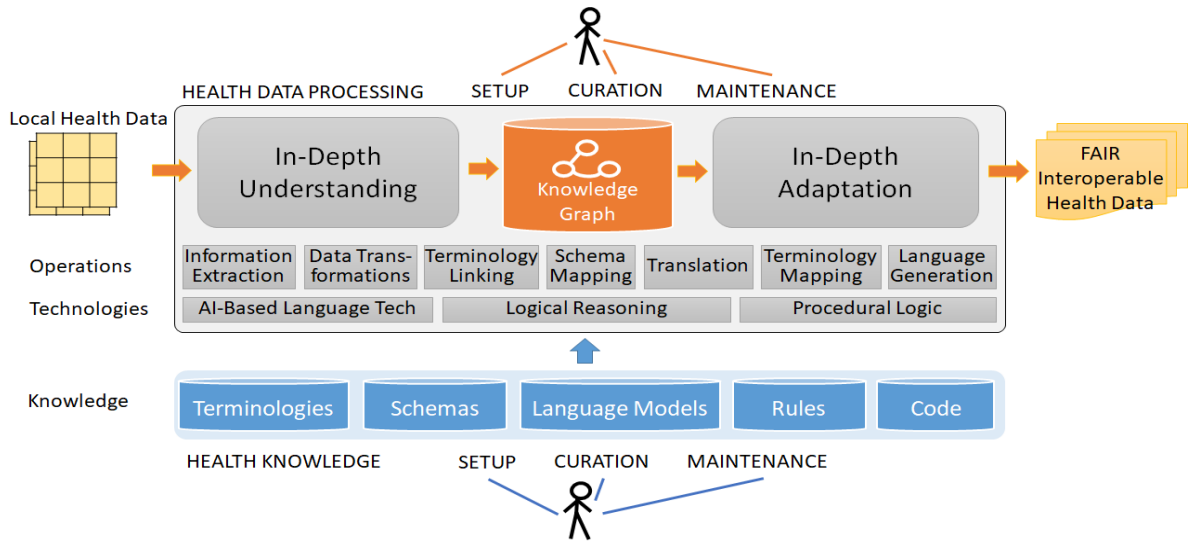


Figure 1: High-level architecture of the Health Services and the way they are overseen by a human data manager (Source: InteropEHRate project)

The two goals are to adapt the system to local needs like data formats, terminologies, practices, and to follow their evolution. To make this process agile, almost all adaptations are

[34] For a detailed description, see https://www.interopehrate.eu/wp-content/uploads/2022/05/InteropEHRate-Third-White-paper.pdf.

carried out through pluggable **knowledge** (depicted in blue at the bottom of Figure 1) rather than undertaking lengthy software development cycles.

4 Tools and methods to facilitate FAIRification of data: an introduction

In addition to the methodology outlined (above) in sub-section 3.5 and Figure 1, the availability of validated knowledge resources and efficient graphical tools is another important technical factor in the FAIRification of data, and hence in the communication between EHR systems and AI-based tools. This section therefore concentrates on the relevant knowledge resources and the various tools which may be used.

4.1 Knowledge tools

With regard to **knowledge**, there are a number of vital resources. One can refer to ontologies, terminologies, classification hierarchies, coding schemes such as SNOMED CT, the International Classification of Diseases (ICD)[35], and other international, widely used, semantic assets[36]. One can also refer to specific rules established by each healthcare organisation to organise clinical and administrative processes and the local codes which support these processes. The selection of validated schemas and language models[37] is also essential to support the automation of processes.

4.2 Tools supporting operations

When looking at relevant **tools**, those tools which support operations and those which are instrumental in creating controlled semi-automatic validation processes should be considered specifically.

In term of operations, **Extract-Transform-Load** (ETL) tools can be mentioned particularly: they are widely available on the market and are increasingly proposed as providing connections to major data-sharing platforms.[38]

Another important operational requirement is the availability of **terminology servers** which allow mapping between different data resources. More and more European Member States have understood the importance of terminology servers not only to enable mapping, translation, and linkage, but also to guarantee the use of validated and updated semantic

---

[35] https://www.who.int/standards/classifications/classification-of-diseases

[36] In many instances, however, these resources are not yet always sufficiently operationalised to be used in all contexts. The problems with operationalisation can include the lack of availability of translation of SNOMED CT into national languages, and the fact that no validated subsets of data are available.

[37] A language model is a type of machine learning model that predicts the next word or sequence of words in a sentence, based on the words that came before it. It learns the statistical patterns of a language from a large dataset of text, allowing it to generate human-like text, translate languages, and answer questions.

[38] This report does not describe ETL tools in detail, although these tools are mentioned in sub-section 5.4 and Annex 4.

resources. National terminology servers have been funded for example in Austria, France, and the Netherlands, to name only a few countries. While not every European country has a single, centrally managed national health terminology server, many countries have implemented systems or tools for managing and distributing standardised healthcare terminologies which are also increasingly FHIR®[39] compliant. Terminology servers have proven to be essential in dealing with the many legacy systems inherited from the past and with national/regional specific issues in a well-governed way. Decisions and investments on terminology servers thus need to take place at different operational levels.

4.3 Tools for semi-automatic validation process

In reference to those tools creating controlled semi-automatic validation processes, one can refer to **AI** and their **learning systems**. Once the data is curated, the automated system (i.e., AI) will be able to learn from the human curation about any errors made (e.g., in information extraction from unstructured text) and gradually improve its predictions. This learning ability, together with the automated replay of pre-recorded human input, guarantees the scalability of human intervention over large datasets and, hence, the capacity to interact with AI-driven decision support tools. This ability is also important for the capacity to use standards in more agile ways. While SNOMED CT is extremely comprehensive as a collection of terms, its use of complex pre-coordinated concepts leads to challenging usability issues. The approach proposed here, which is dependent on learning systems, enables a progressive shift to take place towards a post-coordination of complex concepts (combining different codes while maintaining their relationship).

4.4 The case of OpenEHR

It is also important to mention the approach proposed by **OpenEHR**[40], a non-profit organisation that publishes technical standards for an open, vendor-neutral, EHR platform, in which archetypes help to define precise clinical concepts with a strong emphasis on semantic interoperability by design. Hence, there is a possible important impact on continuity of care, data quality, and on the capacity of the system to interact with AI-driven decision support systems.

Other initiatives have been developed by individual Member States, such as the clinical building blocks[41] developed by the Netherlands, but so far they have had limited results.

4.5 Validated authenticated sources

The governed maintenance of knowledge-based resources (see Figure 1 – above), which is meant to be used by all systems, is an important enabler.

---

[39] https://fhir.org
[40] https://openehr.org
[41] https://zibs.nl/wiki/HCIM_Mainpage

Governed maintenance can also apply to data: this is particularly the case for data related to medicinal products – many AI-driven decision support systems rely on that data. The availability of validated and structured medicinal products datasets can thus be seen as a key contributor to the use of AI-driven decision support tools.

Only a limited number of European Member States have invested in the production of a **validated and structured database** which is then used by all clinical systems operating in the country. Other Member States have preferred to define official reference documents, leaving the market to develop the actual databases.

The progressive and compulsory implementation of the ISO IDMP[42] **standard**, complemented by reference to validated and maintained identifiers by the pharmaceutical industry and by European and national regulatory (marketing authorisation) authorities, brings about a totally new perspective. Structured and coded data can now flow along the whole healthcare value chain and 'cross the border' between regulatory and clinical domains. A European database for drugs and substances has already been validated[43] and will increasingly be used by EHR systems once the core structured data is made publicly available.

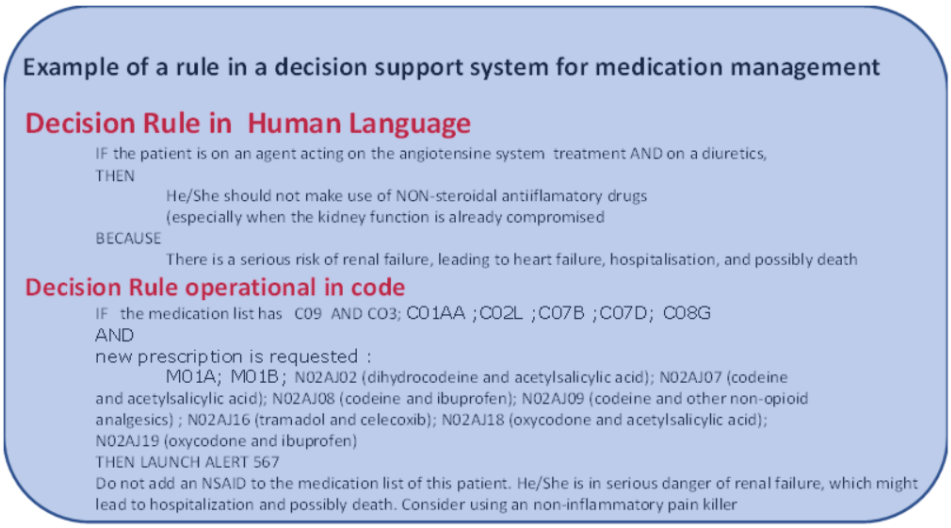


Figure 2: Example of a rule-based decision support facilitated by the use of a validated ISO IDMP compliant medical products dataset[44]

This evolution in medicinal products will have a very important impact not only in terms of global efficiency and pharmacovigilance but also on the actual use of AI-driven decision

[42] https://www.ema.europa.eu/en/human-regulatory-overview/research-development/data-medicines-iso-idmp-standards-overview

[43] The primary EU databases for drug and substance information include the European Medicines Agency (EMA) Substance Registration System (EU-SRS) (https://www.ema.europa.eu/en/human-regulatory-overview/research-development/data-medicines-iso-idmp-standards-overview/substance-product-organisation-referential-spor-master-data/substance-product-data-management-services).

[44] https://unicom-project.eu/

support systems dealing with interactions between the data as well as between that data and other data. It will also reduce drastically the need for data transformation. Medicinal products are central elements in all but two[45] of the information domains named by the EHDS Regulation.

4.6 Competitive advantage of integrated ecosystems

Healthcare organisations that have begun integrating AI-based algorithms into their EHR systems for routine use are currently still limited in number. As described in Annex 2, these integrations aim at improving clinical decision-making, enhancing patient care, and streamlining workflows.

Unsurprisingly, the organisations that have begun this integration are often health management organisations which have fewer interoperability challenges to resolve (since they manage a highly integrated ecosystem which defines its own rules). This is the case, for example, of New York City's Mount Sinai Health System[46] which uses AI tools integrated with EHR for predictive analytics, such as the early detection of sepsis and patient deterioration. Other examples include Kaiser Permanente[47], which focuses on predictive modelling and risk stratification and the Mayo Clinic[48] which employs AI algorithms in its EHR systems for various applications, including clinical alerts and personalised treatment recommendations.

In Europe, various healthcare organisations have moved forward on this front. Examples of organisations/institutions can be cited from Finland, Spain, and the United Kingdom. NHS England has funded and rolled out multiple AI tools into care pathways (retinal/ophthalmology screening, cancer triage pilots, and prediction tools) and supports national health service trusts to integrate AI into clinical workflows/EHRs[49]. Helsinki University Hospital (Finland) has active collaborations that embed diagnostic/predictive AI into hospital data lakes and clinical workflows for rare disease diagnosis and other routine decision support[50]. Hospital Sant Joan de Déu Barcelona (Spain) is running an AI/"Health Copilot" lab with Microsoft to evaluate and integrate AI into its command centre and clinical pathways (i.e., planned/ongoing integration with hospital workflows/EHR-connected systems)[51].

Large companies producing EHR systems with advanced functionalities already propose the smooth integration of AI modules. For example, EPIC[52] focuses on clinical decision support,

[45] The two exceptions are "medical imaging studies and related imaging reports" and "medical test results, including laboratory and other diagnostic results and related reports".
[46] https://www.mountsinai.org
[47] https://healthy.kaiserpermanente.org/front-door
[48] https://www.mayoclinic.org
[49] https://www.england.nhs.uk/long-read/artificial-intelligence-ai-and-machine-learning
[50] https://www.tietoevry.com/en/newsroom/all-news-and-releases/articles/2024/02/ai-powered-healthcare
[51] https://www.sjdhospitalbarcelona.org/en/news/sjd-barcelona-childrens-hospital-microsoft-are-collaborating-creation-ai-laboratory-development-digital-health-projects
[52] https://www.epic.com

population health management, and predictive analytics, while Oracle's Cerner[53] has prioritised features for clinical workflows, including sepsis prediction and adverse event detection. Many organisations which have purchased such products still fail, however, to be able to make use of these AI-based tools (due mainly to difficulties to comply with requirements that users consider to be burdensome, in particular since users are often under very high workload pressure[54]).

Taking this difficulty into account, purely as an example, EPIC proposes two features:

- Ambient listening technology to generate real-time progress notes from conversations held during medical examinations.
- The use of AI to help drafting responses to patient inquiries with an adequate level of empathy (physicians can review and personalise these notes, thereby easing communication and improving responsiveness).

AI is also used to summarise patient histories and highlight what changes have occurred since the patient's last visit, while context-specific summaries and the ability to see source citations helps to increase trust in AI outputs. EPIC, nevertheless, still emphasises the importance of having "a human in the loop".

In Europe, although many proofs of concepts are being developed, very few are actually being deployed. The 2025 study commissioned by the European Commission on the deployment of AI in healthcare [55] confirms this statement:

- Most AI solutions are in **testing, validation, or pilot phases**, often confined to individual hospitals or regions, and only a small number of tools have achieved **full integration into national health systems**,
- Deployment is more advanced in countries with **strong digital health infrastructures** and targeted AI strategies, leading to fragmentation across the EU.

The study also confirms that AI is most commonly used in **medical imaging and diagnostics**, particularly in **radiology, pathology, dermatology, and ophthalmology**. AI tools are increasingly explored in **public health surveillance**, **drug discovery**, and **personalised medicine**, although many remain experimental.

The introduction of AI-driven support tools, however, remains problematic for many organisations operating EHR systems as it often requires substantial improvement and important reengineering changes which may have important investment costs. EHR systems vendors also act as gatekeepers and may be hostile to the introduction of AI tools.

---

[53] https://www.oracle.com/corporate/acquisitions/cerner/
[54] E.g., https://www.forbes.com/sites/edwardsegal/2024/07/23/ai-adds-to-the-workload-and-stress-of-employees-report-says/
[55] European Commission: Directorate-General for Health and Food Safety, EEIG, Open Evidence and PwC, Study on the deployment of AI in healthcare – Final report, Publications Office of the European Union, 2025, https://data.europa.eu/doi/10.2875/2169577.

Furthermore, even well-intentioned solutions cannot be run in settings that lack basic infrastructure.

The deployment of AI tools to support clinical documentation seems to be the currently most active business segment in Europe, and several dynamic SMEs are engaged in it. One may cite, for example in Belgium, solutions such as Tirohealth[56] and Amaron[57] which combine AI-supported data capture at the source with FHIR transformation, and anticipate national and European evolutions. But many others are also available, such as Squire[58], Mindoo[59] or Cavell[60], and are quickly gaining ground, as is shown by the number of hospitals which are adopting those solutions. These solutions already go far beyond the "Voice to Text" function, and are thus both "creating" data and consuming data in value-added processes.

Aside from any technical and interoperability aspects, for advanced AI tools to be integrated into daily routine by EHR system users, they need to comply with a seemingly paradoxical challenge: real success comes only when AI disappears, but – in order to be accepted – AI also needs to be fully explainable. Recent experiments have shown that, although mistrust and request for extended explainability are common in the testing phase of development, clinical users are often disappointed when the service ends.[61]

5 Key barriers to data fluidity identified

To date, **five barriers** to data fluidity have been identified: incomplete semantic standardisation and interoperability of data; some aspects of human-driven coding; the absence of high-quality data; a lack of continuous data governance/iterative design; and a lack of openness to third-party design.

5.1 Incomplete semantic standardisation and interoperability

While the EHDS Regulation pushes interoperability, **semantic standardisation is still incomplete**. Many local systems depend on implicit conventions or partial coding schemes (which are often tied to reimbursement), and not on robust clinical semantics. Initial steps are foreseen by the EHDS Regulation, however, such as the use of codes for primary diagnostics from 2029 onwards.

[56] https://www.tiro.health
[57] https://amaron.be/en/bridging-clinical-documentation-and-fhir-how-tiro-health-and-amaron-are-supporting-the-transformation-of-belgian-healthcare-data/
[58] https://squire.eu/
[59] https://mindoo.ai/
[60] https://www.cavell.ai/en
[61] Evidence from the IDERHA project (https://www.iderha.org/) – Improving clinical decision-making and enhancing patient access to health innovations through better use of health data – communicated by Jens Declerck – Health Data Quality Manager, i~HD.

If they are not based on shared, integrated standards or structured terminologies (like SNOMED CT, LOINC, ICD, FHIR® resources with semantic profiles), algorithm-based tools cannot reliably interpret or process data. Even where these standards exist, operationalisation remains patchy. There can be difficulties with e.g., translations, validated subsets, or alignment with national requirements.

5.2 Heavy reliance on human-driven coding in non-clinical silos

Human-driven coding often occurs in non-clinical settings, and **may currently be relied on heavily**. In many EU countries, substantial financial and human resources are used to code data often in relationship to reimbursement optimisation. The hard coding of data (using structured terminologies) is often handled by administrative teams, whose members may operate in a disconnected way from clinicians. Clinicians often view data entry as a burden that has unclear benefits: this reduces their motivation to ensure the capture of structured, high-quality data that algorithms could later use.

5.3 Risks from algorithm-based tools requiring quality data

AI tools (examples include predictive risk models, Natural Language Processing (NLP), and image analysis) rely on **high-quality, harmonised, context-rich data.** Unfortunately, they rarely draw on such high-quality data.

Local data heterogeneity, a lack of traceable mappings, and evolving clinical language mean that:

- AI tools trained in one context may underperform or mislead in another context,
- Even minor data misinterpretations can have large clinical impacts.

If AI models trained in one institution are deployed elsewhere, model drift and performance degradation occur when data definitions, population characteristics, or workflows differ – reinforcing the need for robust data-quality governance.

5.4 Limited progress on continuous data governance and iterative design

Interoperability is **not a one-shot technical procedure**, but an ongoing process supported by a governance process which combines at least three elements:

- **Automated tools** (e.g., ETL processes, terminology servers, and ontology mappings),
- **Human oversight** (e.g., by knowledge managers/data managers),
- **Iterative design and validation** which adapts to local workflows and data evolution.

Many organisations, however, still lack agile, well-resourced forms of governance to support this iterative lifecycle.

Although implementation success relies on addressing all key factors listed above, clearly defined roles (Chief Information Officer (CIO), Chief Medical Information/Informatics Officer (CMIO), IT lead, clinicians, compliance officers, and vendors have a major influence in reducing implementation failures.

5.5 EHR-centred ecosystems can exclude disruptive external innovation

EHR vendors have a crucial role to play in ensuring the necessary data fluidity to feed AI-CDSS.

Large EHR vendors increasingly offer built-in algorithmic modules for applications like clinical decision support and predictive analytics. However, these capabilities often remain locked in by there being **insufficient local data quality**, a **lack of interoperability**, **or heavy implementation burdens**.

EHR vendors may also act as **gatekeepers** who resist the disruptive integration of external AI tools for various reasons e.g., to preserve their market control or to avoid the costs of major system reengineering. Such an approach can be referred to as an EHR-centred ecosystem or a "closed ecosystem". This kind of approach can slow or stifle innovative algorithm adoption.

This barrier to innovation is compounded by regulatory ambiguities – for example, under the EU Medical Device Regulation (MDR)[62], many algorithm-based tools qualify as medical devices (Class 2b or higher) which adds heavy burdens of compliance for vendors.

The AI Act[63], MDR, and the EHDS will increasingly require interoperable audit logs, standardised APIs and full traceability of data exchanged with AI-CDSS – indirectly limiting vendor closed ecosystem strategies.

6 Linking FAIR data and data fluidity: AI-CDSS models of deployment

The capacity of AI-CDSS systems to provide their expected value is directly related to both the level of FAIRification and the level of fluidity of data in and between systems.

Healthcare organisations may deploy AI clinical decision support systems (AI-CDSS) using three different technical models*:*

- **On-premises systems** tightly integrated into local EHR/PACS (Picture Archiving and Communication Systems) infrastructure. This setup is often prioritised when data localisation or GDPR constraints are strong.

[62] Regulation (EU) 2017/745 (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02017R0745-20250110).
[63] https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

- **Cloud-based systems** delivering scalable AI inference via secure APIs (which is common for imaging, risk scoring, and NLP).
- **Hybrid systems** combining local data retention with cloud-executed models. The implications for data governance and interoperability differ across these models.

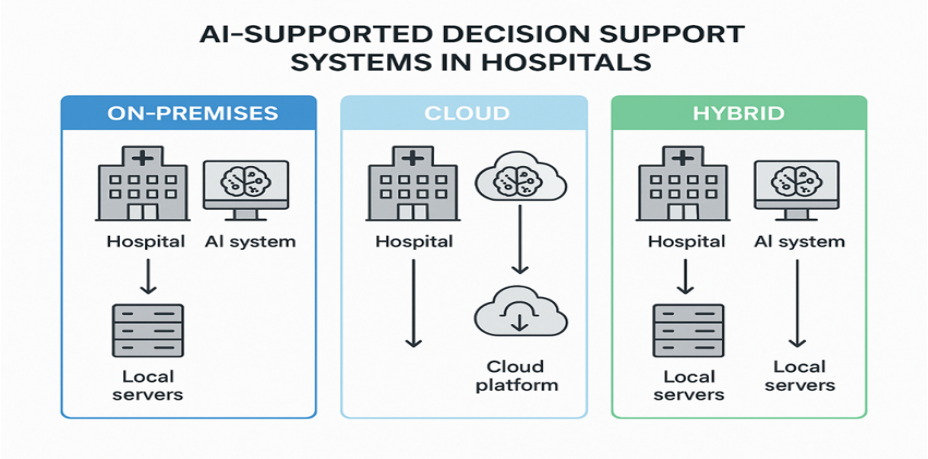


Figure 3: Technical implementation models for AI-DSS in hospitals

AI-DSS used by individual healthcare providers often tend to rely on cloud or hybrid models due to scalability and processing requirements, although on-premises system implementation remains common where privacy, latency, or regulatory constraints apply. The use of cloud or hybrid approaches require an updated and proactive approach to cybersecurity.

Furthermore, usability aspects such as response time also need to be taken into account as the routine use of cloud-based models require advanced connectivity.

7 Functional and Non-Functional[64] Requirements of FAIR Data for AI-(C)DSS

In this section, we present a high-level description of the functional and non-functional FAIR data requirements for AI decision support tools by considering eight essential pillars: interoperability, data quality, workflow integration, user interface and user experience, security and privacy, safety and performance, deployment and adaptation, legal and regulatory compliance and patient-facing DSS.

Most of those pillars have already been mentioned in previous sections of this report. Requirements are context and use-case sensitive; hence, not all requirements apply

[64] Functional requirements define what a system must do (features, actions, business rules), like user login or search; non-functional requirements define how well the system does it (quality, constraints like speed, security, usability, performance), ensuring a good user experience, stability, and efficiency.

everywhere. They are discussed here – in the EU context – through the use of simple statements which summarise trends and/or challenges.

Functional requirements are foundational for the use of AI-DSS. As much as possible, the non-functional requirements are aligned with the data quality assurance perspective, and legal/regulatory references have been added whenever needed. High-level legal and regulatory compliance aspects have also been summarised separately. While requirements might sometimes appear to go beyond the FAIR data requirements, they all have important implications for EHR systems.

7.1 Interoperability and Data Exchange

Functional Requirements	**EU-Specific Nuance**	**Non-Functional Requirements**	**EU-Specific Nuance**[65]
HL7 FHIR/ SMART integration	EU-driven standards are converging toward FHIR® ; national requirements vary (e.g., Germany's MIOs, France's CI-SIS[66])	Compliance to standards	Must comply national eHealth interoperability frameworks and EEHRxF
Multimodal data ingestion	Imaging compliant with DICOM; pathology compliant with DICOM-WSI (Whole Slide Imaging). Increasing demand for standards in new domains (e.g., behavioural and interaction data, sensor and Internet of Things (IoT) data, phenomics, PREMs & PROMs, videos, and fusion layers) In most cases, even if metadata (e.g. DICOM) are present, semantics remain weak or are insufficiently aligned	High-throughput ingestion	Align with EHDS for secondary use

[65] Highlights what is EU-specific in comparison with other situations (such as that of the USA).
[66] https://gnius.esante.gouv.fr/en/regulations/regulation-profiles/sis-interoperability-framework-ci-sis

Functional Requirements	**EU-Specific Nuance**	**Non-Functional Requirements**	**EU-Specific Nuance**[65]
Real-time data eventing	Supports early-warning and deterioration alerts	Low latency	Must handle cross-border data flows under the GDPR with appropriate safeguards
IoT / device integration	Must ensure CE-marked medical devices interoperability	Reliability	Cybersecurity requirements need to be aligned with the NIS2 Directive[67]
Bidirectional AI-EHR communication	Must maintain auditability for MDR high-risk AI	Traceability	Logs must meet AI Act[68] Art. 12 requirements

7.2 Data Quality, Fairness, Bias and Preprocessing

Functional	**EU Emphasis**	**Non-Functional**	**EU Emphasis**
Data completeness rules	Art. 10 EU AI Act[69]: training data must be relevant, representative, and free of errors	Data accuracy	Must follow AI Act Annex IV data documentation requirements
Metadata requirements	Provenance tracking required for AI Act high-risk AI	Consistency	Mandatory under AI Act Art. 10

[67] Network Information Security 2: https://digital-strategy.ec.europa.eu/en/policies/nis2-directive
[68] https://artificialintelligenceact.eu
[69] https://artificialintelligenceact.eu

Functional	EU Emphasis	Non-Functional	EU Emphasis
Bias and fairness monitoring	High-risk AI must undergo bias assessment, mitigation, and documentation	Fairness	Monitoring must show no disproportionate performance across demographic groups (AI Act Art. 9 and risk management obligations)
NLP/Unstructured data pipelines	Must meet health-data processing rules under the GDPR	Privacy	Must implement data minimisation and "privacy by design"

7.3 Workflow Integration and Human-AI Teaming (EU-Specific Governance)

Functional	EU Emphasis	Non-Functional	EU Emphasis
In-clinical workflow AI triggers	Must follow national clinical pathway standards (e.g., NICE in UK[70] or HAS in France)	Seamlessness	Consistent with EU ergonomics and safety norms
Care escalation pathways	Align with EU hospital safety frameworks	Alert optimisation	Must reduce alert fatigue; required under AI Act risk management
Clinician feedback loops	Evaluation data is part of post-market surveillance for high-risk systems	Adoption	Evidence required for MDR and AI Act conformity assessments
Human oversight	Mandatory: AI Act Art. 14 requires meaningful human oversight	Human factors	Must assure that the clinician can interpret and override the AI

[70] https://www.nice.org.uk

7.4 User Interface (UI)/User Experience (UX), Explainability, and Transparency

These conditions have an especially strong emphasis in the EU.

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Explainability	Required for high-risk AI (AI Act: transparency obligations)	Interpretability	Must enable the clinician to understand the model logic sufficiently to override decisions
Uncertainty & risk communication	Supports AI Act's risk management and transparency duties	Readability	Explanations must be understandable by intended users (clinician or patient)
Model provenance	Required by MDR and AI Act (technical documentation)	Documentation	Must follow Annex IV (technical documentation) including training data summary
LLM safety	Large Language Models (LLMs) used in healthcare must follow "General-purpose AI" obligations and high-risk add-ons	Hallucination prevention	Required under AI Act Art. 52 (transparency)

7.5 Security, Privacy, and Data Protection

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
GDPR compliance	Strictest global standard; lawful basis, Digital Privacy Impact Assessment (DPIA) required	Privacy	Must apply "privacy by design" (GDPR Art. 25)

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Encryption, Role-Based Access Control (RBAC)	Must meet NIS2 cybersecurity essential requirements	Cybersecurity	Must ensure CE-marked software meets MDR Annex I safety principles
Consent management	GDPR and EHDS requires explicit consent in certain contexts	Trust	Patient must be informed when interacting with AI (AI Act transparency) See in particular Article 61: Informed consent to participate in testing in real world conditions outside AI regulatory sandboxes
Federated learning	Must maintain data localisation rules in some countries	Confidentiality	Must avoid cross-border data transfers unless GDPR compliant

7.6 Safety, Performance, Monitoring and Post-Market Surveillance

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Clinical validation	Mandatory clinical evidence under MDR Annex XIV	Scientific rigor	Must generate evidence comparable to EU Health Technology Assessment (HTA) standards[71]

[71] https://health.ec.europa.eu/health-technology-assessment/implementation-regulation-health-technology-assessment_en

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Post-market monitoring	Strong obligation under AI Act Art. 61 and MDR ¨Post-Market Surveillance (PMS) plan	Continuous monitoring	Surveillance reports mandatory (high-risk AI)
Incident reporting	Must follow MDR Vigilance + AI Act serious incident reporting	Reliability	Reporting windows strictly defined
Drift detection	Required as part of continual risk management	Performance stability	Documented in technical file updates

7.7 Deployment, Adaptive AI Lifecycle and Change Management

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Change management	Must follow MDR significant change rules (no U.S. Food and Drug Administration (FDA) Predetermined Change Control Plans for Medical Devices (PCCP) equivalent)	Stability	AI Act will require updates to be documented and possibly re-certified
Versioning & rollback	Must preserve traceability for regulators	Maintenance	Audit trail must meet AI Act Art. 12
Cloud/edge deployment	Must ensure compliance with EU data sovereignty expectations	Reliability	EU-hosted clouds or sovereign solutions preferred

Functional	EU-Specific Nuance	Non-Functional	EU-Specific Nuance
Silent testing	Part of conformity assessment and clinical evaluation	Safety	Must demonstrate no harmful effects in real-world test phases

7.8 Patient-Facing Clinical Decision Support (CDS) and Digital Health Tools

Functional	EU Emphasis	Non-Functional	EU Emphasis
Symptom checkers & triage bots	Clearly fall under high-risk or general-purpose AI obligations	Accuracy	Must be validated for clinical safety (MDR)
Remote monitoring CDS	Often Class IIa/IIb under MDR	Reliability	Device-software interoperability must be CE compliant
Patient education	Must follow transparency requirements	Readability	Local language and health literacy requirements (EU multilingual standards)[72]
Patient data rights	GDPR: right of access, to object, to rectification	Trust	Patient must be informed when AI is used (AI Act)

[72] See, for example, the European Health Literacy Project (HLS-EU) which provides definitions, measurement frameworks, and principles highlighting that health literacy includes the ability to access, understand, appraise, and apply information – implying that health information should be accessible across different languages and literacy levels.

7.9 Summary of Regulatory and Compliance requirements

Functional	**EU legal references**
Classification	AI-CDSS for diagnosis/treatment = High-Risk AI under AI Act and is usually CE-marked under MDR Class IIa-IIb
Quality assurance	AI Act Art.10
Technical documentation	Must follow AI Act Annex IV and MDR Annex II
Risk management	Mandated by AI Act Art. 9 (ISO 14971[73] alignment)
Human oversight	Required by AI Act Art. 14
Auditability & logging	Required by AI Act Art. 12
Conformity assessment	Performed by Notified Bodies[74]
Post-market surveillance	Strong obligations under MDR and AI Act Art. 61

8 Use case examples and their FAIR requirements

This section aims at documenting FAIR data requirements adapted to specific contexts and clinical situations.

Two broadly similar use cases (Type 2 Diabetes and hip replacement) were discussed briefly with the participants of the 5 December 2025 workshop (see Annex 4). One of them (Type 2 Diabetes) has been developed in more detail as the requirements for both use cases were roughly similar.

A supplementary use case has been added (on congestive heart failure) to highlight the specificity of the hospital context.

[73] https://www.iso.org/standard/72704.html

[74] Notified Bodies (NBs) are independent, third-party organisations designated by EU countries to assess the conformity of certain risky products (like medical devices or personal protective equipment) with EU regulations before they can be sold in Europe.

These use cases do not, of course, cover the full spectrum of all the FAIR requirements, but they are meant to illustrate in a more concrete way the high-level requirements presented in the various tables in Section 7 of this report.

8.1 Use case 1: EHR-AI interaction for congestive heart failure

This scenario takes place in a hospital setting and involves different systems, namely a hospital EHR, a patient-facing APP and connected devices and an AI-CDSS.

A 74-year-old male patient is affected by chronic heart failure (CHF) with a reduced ejection fraction (HFrEF, EF 30%), hypertension and atrial fibrillation, and with mild cognitive impairment. He is using a remote monitoring scale and a blood pressure cuff. He was admitted to hospital for an acute CHF episode two times last year. He feels “a bit more tired than usual,” but this is not considered severe enough to call his cardiologist. His weight has increased by 1.7 kg over 72 hours, his resting heart rate has risen, and device data shows decreasing nighttime variability. None of these data are cross-checked manually by the medical staff. The cardiologist following this patient manages many patients with remote monitoring devices; she relies on seeing their EHR trends, but is overwhelmed by notifications and wants AI to filter out background noise but not to take over her clinical judgement. The AI-CDSS main goals are early detection of decompensation, avoidance of admission, and discharge optimisation.

The data workflow is shown immediately below and is explained with a more detailed description of requirement for each level (see Levels 1 to 7 which follow).

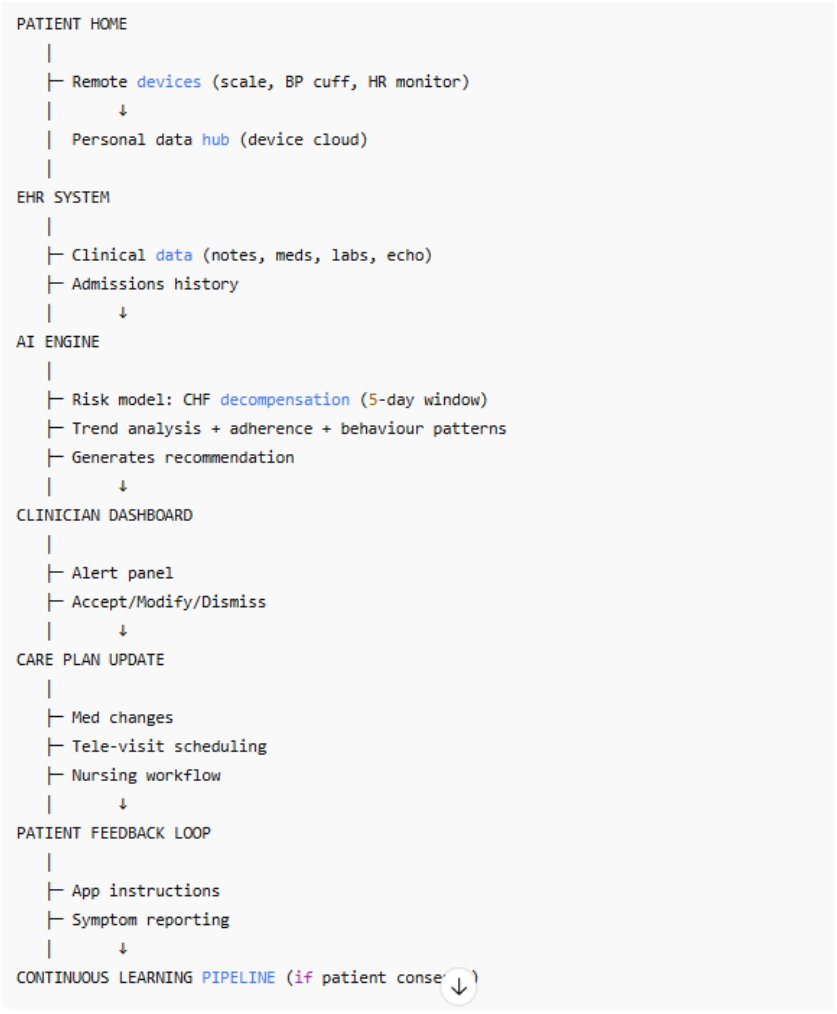


Figure 4: EHR-AI Interaction for Congestive Heart Failure data workflow

Level 1a: **Patient and Data Sources**

Data Sources	**Standards**	**FHIR® Resource**
• Home weight scale • Blood Pressure (BP) / Heart Rate (HR) monitor • Symptoms (dyspnoea, oedema • Hospital vital signs and labs	• IEEE 11073 (medical device data) • ISO/EN device certification • Timestamp and provenance metadata	FHIR® DeviceObservation (API Push)

Level 1b: **Personal / Remote Data Hub** (Remote monitoring- Device vendor cloud / Hospital RPM system)

Functions	**Standards & API**	
• Device data normalisation	• FHIR® R4 Observation	

Functions	Standards & API	
• Quality checks (missing data, outliers) • Patient identity matching	• OAuth2 / OpenID Connect • RESTful APIs **FHIR® API → Hospital EHR**	

Level 2: **Hospital EHR Core System** (Source of trust)

Data objects	APIs	
• Labs (BNP, creatinine) → **LOINC** • Diagnoses → **SNOMED CT** • Medications → **IDMP, UCUM, EDQM** • Echo reports → **DICOM** • Admissions history → **FHIR Encounter**	• FHIR® REST API (Read) • HL7 v2 (legacy lab feeds) **FHIR® Bulk Data / SMART-on-FHIR®**	

Level 3: **AI Decision Support Engine** (Not autonomous – advisory only)

AI Inputs	Outputs	Standards
• Time-series vitals • Weight trends • Prior admissions • Medication adherence	• CHF decompensation risk score (5-day window) • Confidence score • Explainability artefacts	• FHIR® RiskAssessment • FHIR® DetectedIssue • Model metadata (provenance, version) **FHIR® CDS Hooks / SMART Callback**

Level 4: Clinician Interaction Layer (Dashboard inside EHR User Interface).

Functions	Actions	Standards
• Non-intrusive alert panel • Trend visualisation • Explainability ("why this alert?")	• Accept • Modify • Override (with justification)	• FHIR® CarePlan • FHIR® ServiceRequest • Audit logs (IHE ATNA) ⬇ **FHIR® Write-back**

Level 5: Care Plan Execution

Functions	Standards	
• Medication orders (CPOE) • Nursing workflows	• HL7 FHIR • HL7 v2 (orders, pharmacy)	

Functions	Standards	
• Tele-visit scheduling • Discharge planning tools	• Scheduling APIs (FHIR® Appointment) ⬇ **Patient-facing APIs**	

Level 6: Patient Feedback Loop

Functions	Standards	
• Instructions and education • Daily symptom reporting • Weight and vital signs confirmation	• FHIR® QuestionnaireResponse • Secure messaging (FHIR® Communication) • Consent management ⬇ **FHIR® Observation / QuestionnaireResponse**	

Level 7: Continuous Learning & Governance (Optional, Consent-Based)

Functions	Standards	
• Data anonymisation • Bias and drift monitoring • Model retraining	• FHIR® Bulk Export • OMOP CDM (optional) • Governance metadata (EU AI Act readiness) 🔁 **Feedback → AI Engine (new model version)**	

The following table offers a perspective of the involvement of human beings in the use case by summarising the system requirements and roles of both the clinician and patient for each FAIR requirement.

FAIR Principle	System Requirement	Clinician Role	Patient Role
Findable	Indexed CHF data (BNP, weight, echo)	Structured coding	Accurate history
Accessible	Secure real-time access	Ethical use, audit	Consent & awareness

FAIR Principle	System Requirement	Clinician Role	Patient Role
Interoperable	Standards-based data flow	Structured entry	Certified devices
Reusable	High-quality longitudinal data	Contextual validation	Lifestyle annotation

Table 1: Fair principles, system requirement and human roles (Congestive Heart Failure use case)

The sequence of the scenario is as follows:

- The AI-CDSS integrates together daily scale data, home BP & HR readings, prior admission history, sodium refill pattern, the last echocardiogram, and loop diuretic adherence.
- An alert is produced: "High Risk (83%) of CHF Decompensation Within 5 Days", a diagnosis which is supported by the pattern of rapid weight gain, rising HR, low adherence, and prior examples of exacerbation.
- The AI Recommendation Panel in the dashboard (FHIR® RiskAssessment and ServiceRequest) shows the following recommendations: increase furosemide by 20 mg for three days, schedule tele-visit within 48 hours, activate nurse follow-up workflow, and send patient instructions: "Monitor weight daily and report breathlessness".
- The cardiologist is proposed by the system to select one of the following options: Accept | Modify | Dismiss | Show Evidence. The AI level of "confidence" is 81%, with 20 similar patients having avoided hospitalisation through early diuretic adjustment.

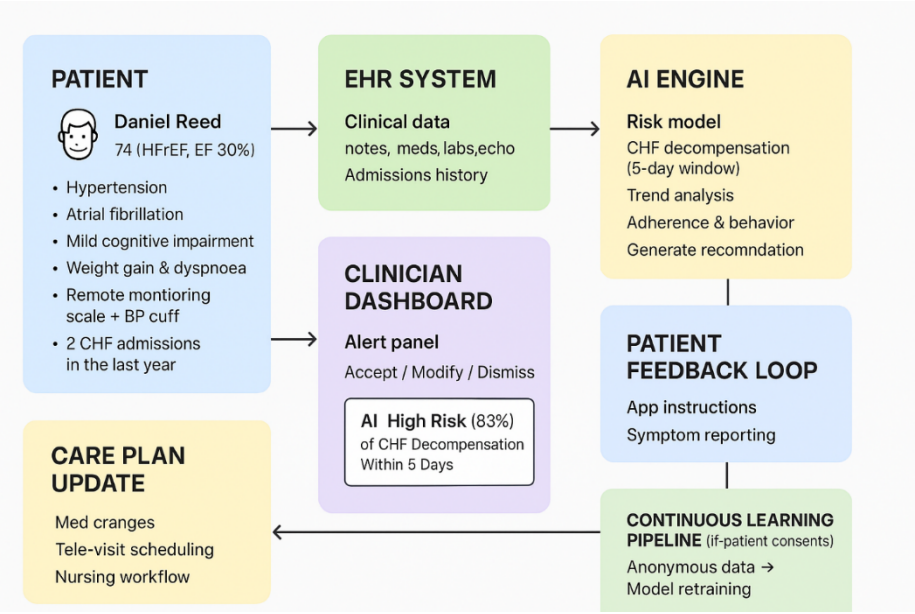


Figure 5: EHR-AI Interaction for Congestive Heart Failure scenario

- The cardiologist remembers that the patient has mild cognitive impairment and that, as a consequence, adherence variations might explain the weight gain. She wants more context before changing the medication schedule.
- She clicks “Show Evidence” and the system displays: Weight trend +2.3 kg since Monday - HR +8 bpm - missed three morning doses - 78% adherence to fluid restrictions.
- She then chooses “Modify Recommendation” as follows : Televisit tomorrow - Nurse call today - Postpone diuretic increase pending evaluation.
- The patient receives an App notification request that confirms the symptoms and is asked to answer a few questions: “Increased swelling? Breathlessness during sleep? Missed medicines?”. He reports mild ankle swelling which triggers AI upgrades to the monitoring level.

The table below details the most important requirements in relation to data capture and sharing.

Findable	Accessible	Interoperable	Reusable
AI and clinicians must rapidly discover **all clinically relevant CHF signals**, across acute and longitudinal timelines.	Relevant data must be available **in real time**, securely, to authorised actors and systems.	CHF data must flow **seamlessly across hospital systems, devices and AI tools**.	Data should support **longitudinal CHF management, audit, and AI improvement**.
Persistent patient identifier across: • Hospital EHR • Remote monitoring platforms • Previous admissions	**Role-based access control** for: • Cardiologists • Ward nurses • Pharmacists • AI decision support system	Adoption of standards: • **HL7 FHIR®** – EHR ↔ AI • **LOINC** – BNP, creatinine, electrolytes • **SNOMED CT** – diagnoses, symptoms • **DICOM** – echocardiography • **IEEE 11073** – remote monitoring devices	**Provenance tracking**: • Who entered data • Device used • Clinical context
Standardised metadata for: • Daily weight (measurement method, time) • BNP / NT-proBNP • Echocardiography (EF, LV size)	**Secure APIs** for: • Real-time weight and vital signs ingestion • Lab result streaming	**Harmonised units**: • Weight (kg) • Fluid balance (ml/day) • BP (mmHg)	**Data quality** indicators: • Missing values • Delayed measurements • Device reliability flags

Findable	Accessible	Interoperable	Reusable
• Diuretic dosing and timing • Fluid balance (I/O)			
Indexed access to: • Prior CHF admissions • NYHA class history • Medication changes • Discharge summaries	Full **audit trail** of: • Data accessed by AI • AI recommendations shown • Clinicians accept/override actions	AI outputs **returned** as: • FHIR® RiskAssessment • FHIR® CarePlan or ServiceRequest	**Contextual** annotations: • Non-adherence episodes • Diet/fluid excess • Acute intercurrent illness
	Bedside and clinician-**dashboard access** without workflow disruption		Clear **consent for secondary use** (no opt-out): • Quality improvement • Model retraining • Population risk stratification
Clinician should use **structured fields** for: • HF type (HFrEF / HFpEF) • NYHA class • Reason for admission • Avoid burying critical CHF signals in free text • Validate completeness of admission and discharge data	**Clinician should** • Access AI insights within authorised clinical scope • Ensure decisions are documented (especially overrides) • Report missing or delayed data access issues	**Clinician should** • Enter data using standard terminologies • Validate device-to-EHR integrations • Avoid non-standard abbreviations and local codes	**Clinician should** • Add clinical context explaining anomalies • Validate trends before escalation • Support safe reuse for learning health systems
Patient should • Ensure remote-monitoring data is correctly linked to identity • Report symptom onset timing accurately • Confirm historical admissions and treatments	**Patient should** Grant consent for: • Use of remote monitoring data • AI-supported clinical decisions Understand how data is used during admission and follow-up	**Patient should** • Use certified, interoperable home-monitoring devices • Avoid manual editing of device data • Report technical problems early	**Patient should** • Provide lifestyle and symptom context • Participate in longitudinal follow-up • Decide whether data can be reused beyond care (no opt-out)

In conclusion, for this scenario to be effective, the following conditions must be met:

- Rapid discovery of early decompensation signals,
- Real-time, secure access during ward rounds,
- Seamless integration of remote and hospital data,
- Clinician trust through explainability and override control,
- Patient engagement as an active data contributor.

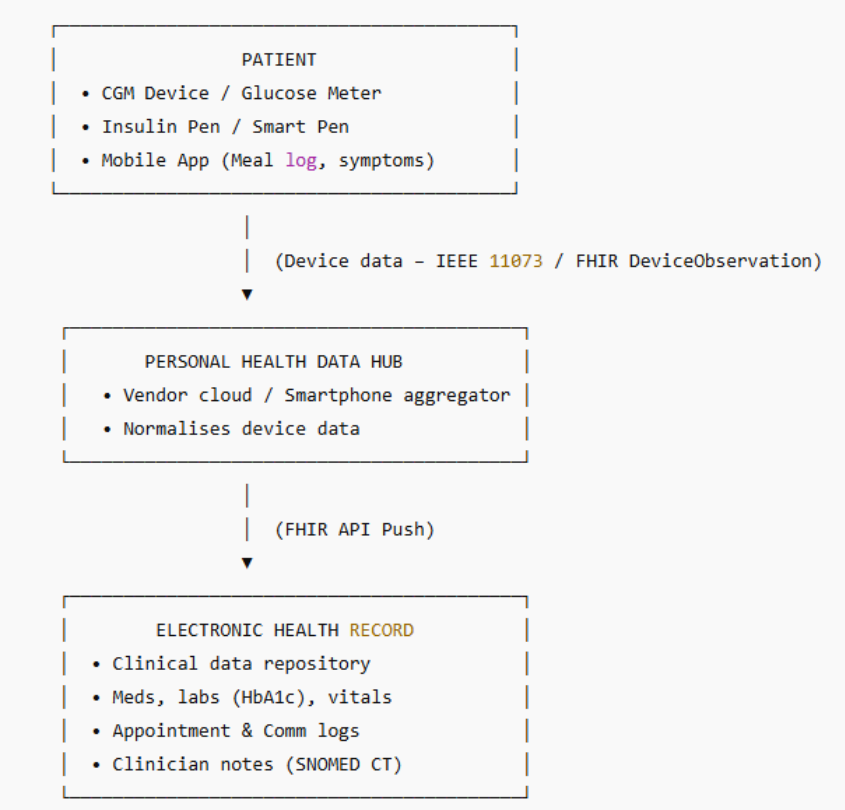
8.2 Use case 2: AI-enhanced management of Type 2 Diabetes in a primary care setting

This scenario highlights the interaction between the following systems: The EHR system used by the general practitioner (GP), the AI-CDSS interfaced with the EHR system of the GP, and the patient-facing Health-APP (and associated connected device) used by the patient.

The patient has Type 2 Diabetes which was diagnosed 10 years ago, is on insulin and metformin, and lives alone. The person struggles with adherence demonstrated by suboptimal control of blood sugars (HbA1c 8.9%).

The GP has a heavy workload and is concerned about alert fatigue and accountability.

The possible data workflow is described briefly below and includes the minimal FAIR data requirements based on current FHIR® standards. It limits itself to identifying the FHIR® resources and the main semantic resources, but access to additional resources might be necessary.



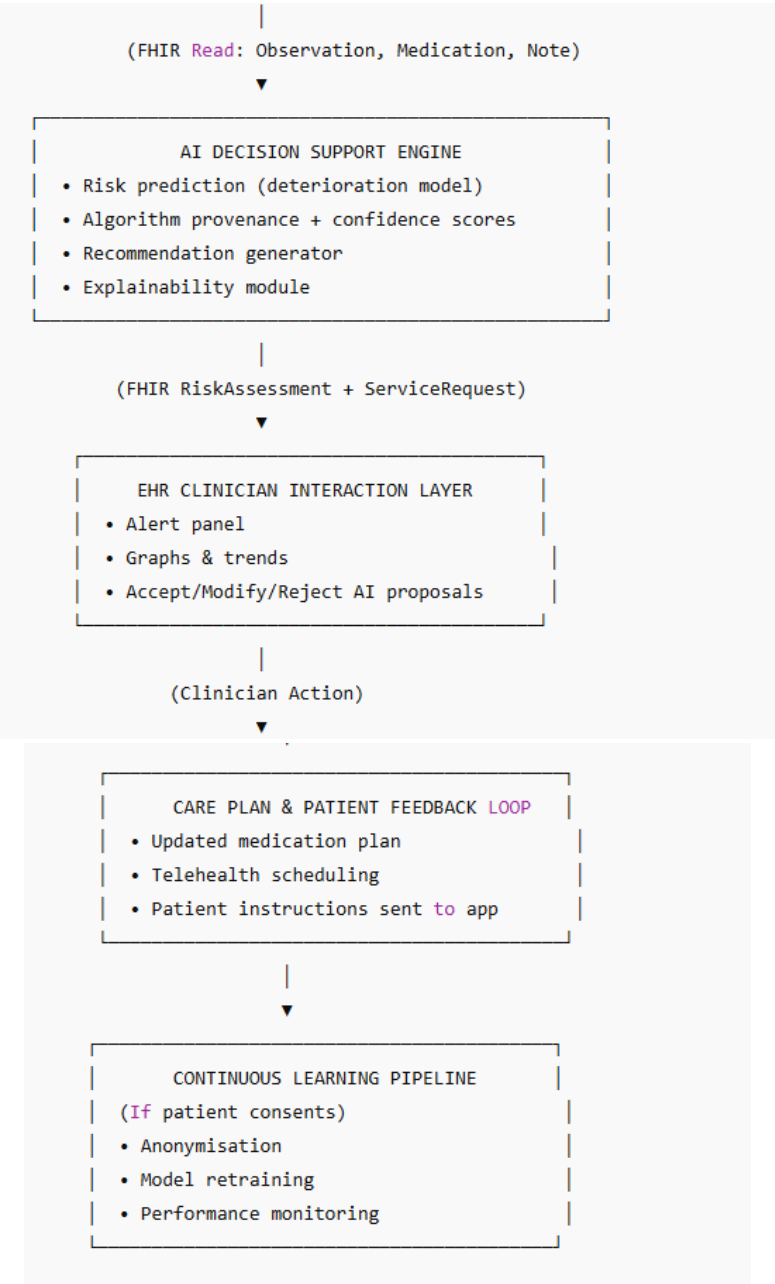


Figure 6: AI- enhanced management of Type 2 Diabetes in a primary care setting data workflow

The scenario postulates a continuous data ingestion pipeline which combines data provided by the patient via her own health app and connected devices and the EHR system.

The AI system has flagged a moderate risk of deterioration and suggests (in a non-intrusive way) increasing the insulin dosage and several other actions.

The clinician may then accept, modify or dismiss the recommendation. He may also request to view the evidence behind the recommended actions.

The evidence might detail the data sources (e.g., literature and datasets) used to train the algorithm and the confidence score.

The decision taken by the GP together with the patient then feeds into an adaptive feedback loop which contributes to the algorithm improvement while the clinician may also adjust his level of interactivity with the AI-CDSS.

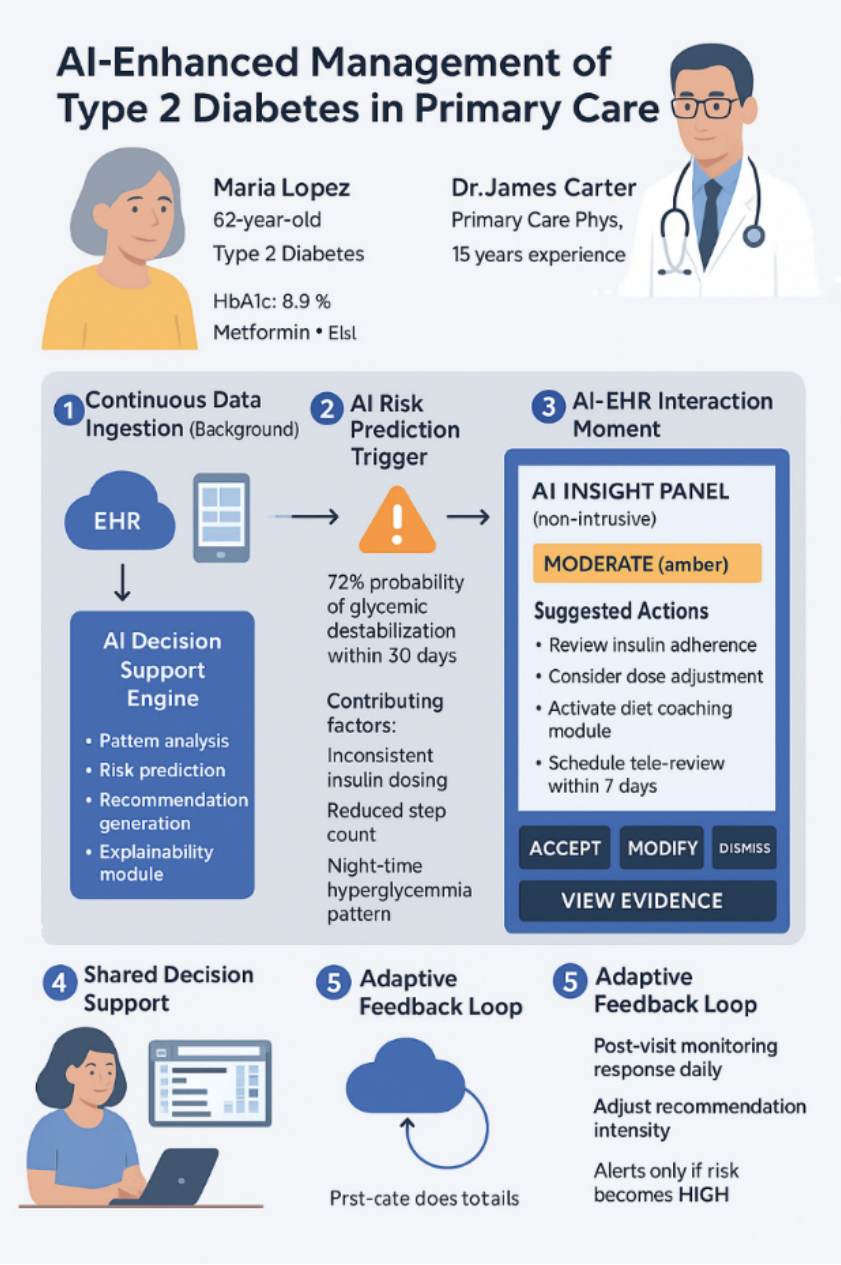


Figure 7: AI-enhanced management of Type 2 Diabetes in a primary care setting scenario

Figure 8 below shows a view of the different modules which need to interact (while the table that follow it offers details of the most important requirements in relation to data capture and sharing).

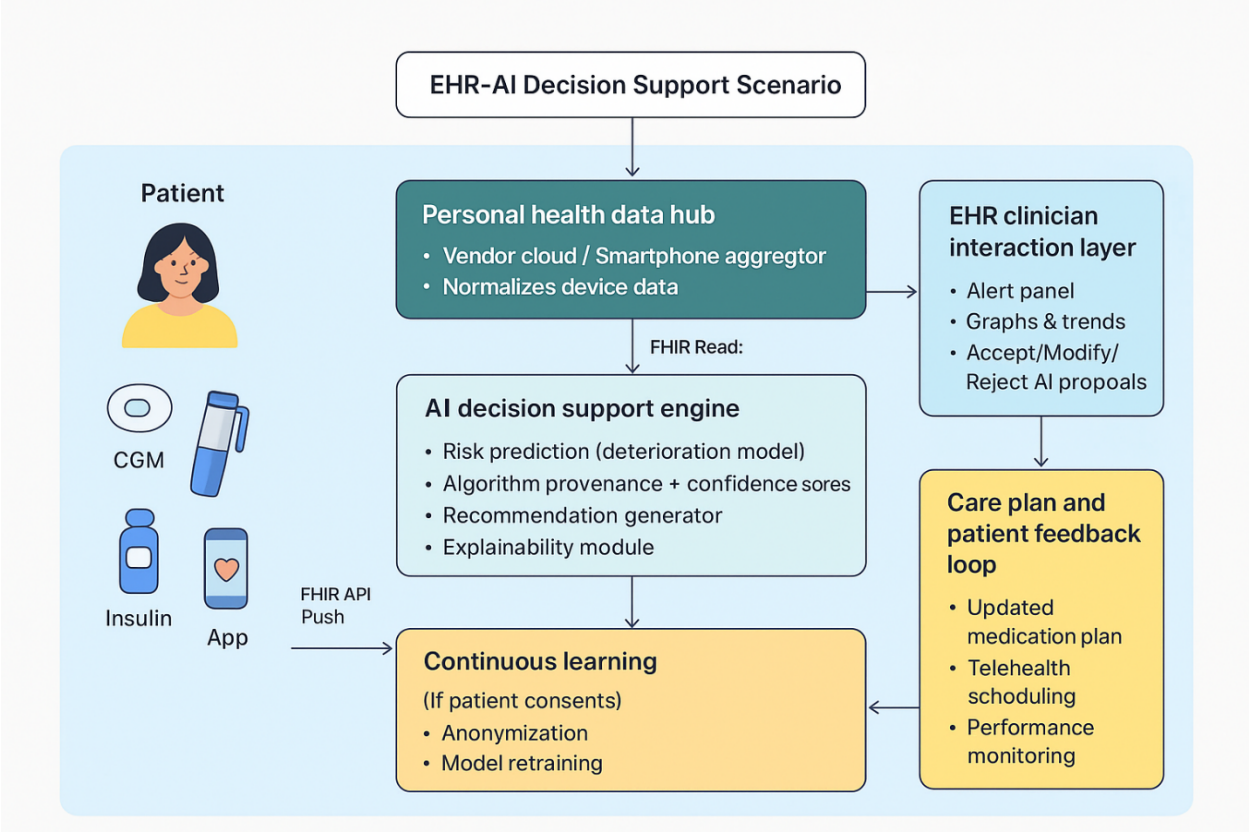


Figure 8: AI-enhanced management of Type 2 Diabetes in a primary care setting - key modules and functions

Findable	Accessible	Interoperable	Reusable
Each data element (glucose logs, insulin dose, HbA1c, meal logs, meds) must have unique, persistent identifiers.	Secure APIs allow the AI system to access EHR data in real time.	• **FHIR®** for clinical data exchange **HL7** for messaging events • **OpenEHR archetypes** for structured clinical concepts • **IEEE 11073** for device data	Data quality thresholds defined (e.g., minimum number of daily glucose readings). Data quality metrics stored with data.
Metadata should describe: • measurement method (e.g., CGM versus finger-stick) • time of capture (timestamp, timezone) • data source: (CGM, EHR, pharmacy) • device source (glucometer model, insulin pen)	• Role-based access control to ensure AI uses only necessary data. • Clear audit trails for data access and use consent management mechanisms.	• European controlled vocabularies (SNOMED CT, ICD, LOINC) • Harmonised data formats for: blood glucose values, Medication units and Time-series data	Provenance capturing: • who entered the data • device calibration status • context of measurement

Findable	Accessible	Interoperable	Reusable
• context (fasting, post-prandial, "skipped meal")			
Data stored in searchable Registries or clinical indexes. Indexed datasets for: • Glucose trends • Medication history • Adherence data	Patients should be able to view AI interpretation (transparency requirement).	AI system must output recommendations in a format the EHR understands (FHIR® ServiceRequest, Observation, RiskAssessment)	Licensing/consent for secondary use in improving AI algorithms must be explicit and auditable.
	Data available even if devices temporarily offline (synchronisation queues).		
Clinician (GP) ensures structured documentation (SNOMED CT, LOINC, EDQM) with the support of appropriate (possibly AI supported) tools and tags key clinical insights for AI visibility.	**Clinician (GP)** checks that appropriate data-sharing consents are activated. Accesses and interprets AI only within clinical context Reports inappropriate or missing data access.	**Clinician (GP)** confirms that data formats and codes support clinical meaning (e.g., "basal insulin increase" represented as a structured order). Avoids unsupported abbreviations or informal coding and validates integration outputs.	**Clinician (GP)** ensures that the notes are clear enough to be reused as training data or as part of longitudinal records. Adds contextual notes improving interpretation. Confirms data validity before action. Supports data reuse for population insights.
Patient ensures personal health devices (CGM, app) are correctly connected and sending identifiable data, uses them consistently, and provides necessary consents.	**Patient** grants data-sharing permission for personal devices/apps and can revoke the permission at any time. Understands how data is shared and why.	**Patient** ensures devices support interoperable export formats; updates apps/firmware when needed. Avoids manually altering data streams and reports technical inconsistencies.	**Patient** specifies whether their data may be reused for AI improvements or only for care delivery. Adds contextual notes improving interpretation. Confirms data validity before action. Supports data reuse for population insights.

The scenario described here continues to position the GP's EHR system in a central position. It is, however, also possible that for such a use case the central position might increasingly shift towards the Personal Health Record (PHR) being under the control of the patient,

although this would have obvious consequences in term of a level of higher MDR compliance[75].

In conclusion, for this scenario to be effective, the following conditions must be met:

- Data must be FAIR-compliant across all systems,
- Clinician judgement must contextualise AI output,
- Patient must be empowered as a data co-producer,
- AI transparency must be maintained.

9 Actors involved in EHR-AI/CDSS interactions

In the previous sections, and in particular in section 8, the main actors involved have only been mentioned in a very generic manner (patient, clinician, EHR, and AI systems developers). However, and as briefly mentioned in section 3.5, the involvement of a large spectrum of people is needed in order to meet the minimum FAIR requirements to interact with AI-CDSS.

9.1 Typology of actors

In this sub-section, we propose a more extensive typology based on existing literature which summarises the key roles of each actor. Depending on the context, the actors involved may vary. We refer to the work of a team of Harvard University which plans to propose an adapted factsheet for each main category of stakeholder.

- **Patients (Persons Receiving Care)**

Patients are an **active data co-producer and decision participant**. They provide accurate symptom and lifestyle information, ensure correct use of monitoring devices, and grant informed consent for data use and AI-supported decision-making. Patients have the right to understand when AI is used, to receive explanations in an accessible language, and to challenge or question AI-supported recommendations. Their engagement directly affects the data quality, model reliability, and clinical safety.

- **Clinicians (End-Users/Decision-Makers)**

Clinicians, including GPs, are the **legal and ethical decision-maker** responsible for patient care. They interpret AI outputs in a clinical context, validate recommendations against patient-specific factors, and decide whether to accept, modify, or override AI suggestions. Clinicians must understand the limitations of AI, document any reasoning when deviating from AI-based advice, and ensure that AI supports – but never replaces – their professional judgement.

[75] This implies at a minimum a class 2a Medical Device.

- **Chief Information Officers (CIOs)**

CIOs are accountable for the **technical integrity, security, and availability** of the digital infrastructure. They ensure that EHR systems, APIs, and integration support FAIR data principles, AI systems are reliably deployed, and cybersecurity risks are managed. A CIO ensures that the organisation can sustain AI technically as a clinical capability, and not simply as a pilot experiment.

- **Chief Medical Information Officers (CMIOs)**

CMIOs bridge clinical practice and digital systems. They are accountable for ensuring that AI outputs align with clinical workflows, guidelines, and patient safety standards. A CMIO ensures clinicians understand the limitations of AI, supports structured clinical documentation, and acts as the clinical owner of decision-support behaviour in the EHR system.

- **EHR Integration Leads (ITL-IT)**

ITL-ITs are responsible for the **practical integration** of AI into the EHR environment. They implement standards-based interoperability (e.g., FHIR®, HL7, SNOMED CT), ensure data flows reliably between systems, and maintain the auditability of AI interactions. Their work determines whether AI is usable at the point of care or remains a technical artifact.

- **Clinical Leads (Physician/ Nursing Leadership)**

Clinical Leads represent the people involved in **frontline clinical reality**. They validate that AI recommendations make sense in real-world practice, define escalation and override rules, and champion safe adoption among clinicians. They ensure that human judgement remains central and that AI supports – rather than distorts – clinical decision-making.

- **AI Safety Leads/Clinical Data Scientists**

AI Safety Leads are responsible for **model safety, performance, and trustworthiness**. They monitor bias, drift, and unintended effects, ensure explainability is clinically meaningful, and validate AI outputs against patient outcomes. They are the technical conscience of the system and a key requirement under the EU AI Act for high-risk AI.

- **EHR / Clinical App Developers**

EHR or clinical app developers design the **user interface, workflow logic, and interaction patterns** through which clinicians and patients experience AI. They are responsible for usability, explainability presentation, alert design, and safe human-AI interaction. They **do not build the AI model**, but they strongly influence how AI impacts clinical decisions. Under the AI Act, poor user interface (UI) design can create foreseeable misuse.

- **AI Vendor / Manufacturers**

AI vendors develop and maintain the AI system as a **regulated medical product**. They provide validated models, documentation, risk management files, and post-market monitoring support. Vendors must ensure transparency, version control, and compliance with MDR and AI Act requirements, but they **do not make clinical decisions**.

- **Data Protection Officer (DPO)**

DPOs safeguard the patient's **fundamental data protection rights**. They ensure lawful data processing, oversee consent and DPIAs, and verify data minimisation and security measures. DPOs ensure that AI innovation does not come at the expense of a failure of patient privacy or lack of GDPR compliance.

- **Compliance Officer (MDR and AI Act)**

Compliance Officers ensure organisational adherence to **medical device and AI regulations**. They translate regulatory requirements into operational controls, verify documentation, and coordinate audits and incident reporting. They ensure that the hospital can **demonstrate compliance** – not just claim to do so.

- **Data Governance Board (DGB)**

The Data Governance Board provides **ethical and strategic oversight**. It arbitrates conflicts between innovation, safety, and compliance; approves high-risk AI deployments; and ensures alignment with institutional values. The board is where technical, clinical, and societal considerations meet.

- **National Competent Authority (NCA)**

The National Competent Authority acts as the external regulator. It oversees compliance with MDR and the AI Act, investigates incidents, and enforces corrective actions where needed. While not involved in daily operations, the NCA shapes accountability and ensures public trust in AI-supported healthcare systems.

9.2 AI Literacy Framework and Healthcare Providers Factsheet

The AI Act describes actors, responsibilities, and liability in a broad perspective since the act does not only consider AI professionals, but also individuals who are interacting only occasionally with the system as well as external consultants, temporary contractors, and third-party vendors. The AI Act covers all AI systems whatever the level of risk implied.

- Art. 3(56) of the AI Act defines "AI literacy" as "skills, knowledge and understanding that allow providers, deployers and affected persons, taking into account their respective rights and obligations in the context of this Regulation, to make an informed

deployment of AI systems, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause."

- Art. 4 of the AI Act requires that: "Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used."

Ensuring a sufficient, adequate and appropriate AI literacy is not an easy requirement of the AI Act to fulfil. Hence, there is a distinct need to operationalise AI literacy through a modular and stratified healthcare **AI Literacy Framework and Healthcare Providers Factsheet**.

Core modules of such a framework/factsheet, which would focus on the AI background, capabilities and limitations, an overview of trustworthy AI, and general best practices, would be applicable to all stakeholders. Meanwhile, specialised tracks will be developed which concentrate on three main groups: Clinicians and patients (interpretive and domain-specific skills); administrative and /legal staff (skills related to regulation, procurement, governance); and IT staff (on system integration, bias mitigation).

Today, this framework/factsheet work is still under development by the Harvard University team, but we provide below[76] an illustration of such a factsheet for clinicians operating in a hospital.

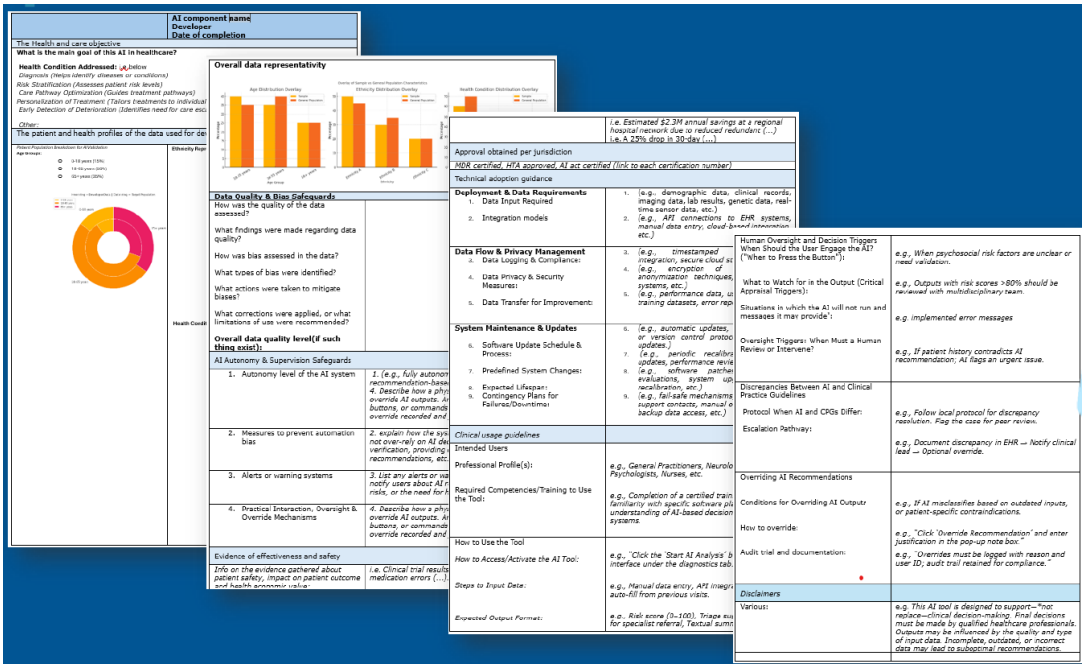


Figure 9: Example of a factsheet in development for a Hospital EHR user interacting with AI-supported algorithms

[76] Illustration provided courtesy of Sofia Palmeri, Ghent University (Belgium) and I~HD https://www.i-hd.eu/team/sofia-palmieri.

The factsheet contains the following elements: the health and care objective, data used for development and validation, data quality and bias safeguards, AI autonomy and supervision safeguards, evidence of effectiveness and safety, approval obtained for each jurisdiction, technical adoption guidance, clinical usage guidelines, and a disclaimer.

10 Priority deployment areas and datasets for AI-CDSS

The effective and meaningful deployment of AI-CDSS has multiple requirements, but the availability of quality data is a necessary preliminary condition – both at the point of care and from reference datasets.

Several earlier sections of this report transmit important messages regarding the quality and quantity of data needed for AI-CDSS, especially sections 3, 4, and 6.

Going beyond the products themselves, as outlined in especially sub-section 3.5, the strategies and measures implemented by healthcare organisations are of prime importance. As also briefly introduced in section 6, a section dedicated to AI-CDSS, the capacity to produce "on-premises" the quantity of data necessary to train the algorithms directly influence the model of deployment.

Hybrid or cloud models of AI-CDSS usually imply additional interoperability requirements. The models developed by Health Management Organisations – as outlined in section 4 – cannot thus be directly used as an implementation reference in a European context. In European Beveridgean systems, and even in European Bismarckian systems, there is, however, currently a clear trend towards regional or sub-regional consolidation leading to more integrated IT ecosystems, at least at hospital level.

Access to datasets via the infrastructure of Healthdata@eu[77] will certainly provide an important contribution, although datasets which have been created for a non-clinical purpose (e.g., reimbursement) are not always trusted by clinicians; hence, the importance of metadata related to (clinical) context. Furthermore, for applications which directly support precision medicine, and to reduce the risk of bias, reliance on local data might be preferable.

Making progress in structuring clinical documentation has been slow, even in organisations which possess several advanced EHR systems. For CDSS which require access to clinical documentation related data, AI tools which capture and structure clinical information might help create a significant breakthrough. The market is increasingly understanding this opportunity, and starting to propose data capture modules to support added value services. The compulsory implementation of the EEHRxF – even when done in phases – will provide an important impetus to the market.

[77] https://acceptance.data.health.europa.eu/?locale=en

In the short term, medicinal data, laboratory results, and imaging present the best chance to meet the FAIR data requirements expected by AI-CDSS. Even if not yet always fully mature or compliant with upcoming technical and semantic standards, those three types of data are usually well structured and include the necessary metadata for adequate data processing. Some European Member States such as Belgium have developed a reference compulsory dataset for medicinal data which contains all necessary identifiers. Furthermore, these reference datasets usually present "inherent" quality properties due to their direct relationship with physical objective entities. For developers of AI-CDSS, targeting these types of datasets means obtaining access to a wider market segment.

The level of risk that AI-CDSS is facing is another key parameter for deployment to occur when based on FAIR data requirements and on quality control constraints which may be adapted to each level of risk.

The third recommendation in section 12 (sub-section 12.3) therefore proposes a three-level Tiered Data Verification Framework which should enable verification that a current level of data FAIRification is compatible with the use of the proposed algorithm. Lower risk levels of AI-DSS, which are targeting clinical knowledge consolidation (i.e., clinical dashboards) and administrative or logistical processes, will logically find a quicker way to market and to integration with EHR systems.

11 FAIR synthetic data versus patient data

FAIR synthetic data is artificially generated data that mimics the statistical properties and patterns of real-world data but contains no actual, real-life.

Using FAIR synthetic data can provide a critical foundation for building trustworthy, scalable, and high-performance algorithms. Soon, a large part of all AI training data will be artificially generated, shifting the development paradigm from reactive (collecting existing data) to proactive (designing ideal data)[78].

As recommended by the 5 December 2025 workshop attendees, this section reflects briefly on the use of FAIR synthetic data to interact with or feed algorithms, taking into account the data contained in a patient summary.

Depending on the technical deployment model of AI-CDSS selected (see section 6), the use of synthetic data might indeed play a crucial role.

In the healthcare domain, synthetic patient summary data must be generated in a way that preserves clinical realism while fully eliminating the risk of patient re-identification. Patient summaries typically combine structured clinical information – such as demographics,

[78] See e.g., https://andrewjpyle.com/blog/synthetic-dataset.

diagnoses, medications, laboratory results, and care episodes – with narrative elements such as discharge summaries or problem descriptions. Because these data types are derived from a specific patient's sensitive health information, the process of data generation must be designed from the outset to ensure privacy, traceability, and reusability.

The creation of synthetic patient summaries begins with the definition of a standardised clinical data model, most commonly based on FHIR® resources or the Common Data Model (see sections 3 and 8 especially). Using a standard model ensures interoperability and supports FAIR principles by making the resulting dataset easier to integrate with external tools and algorithms.

Thus, it is clear that **the implementation of FAIR principles for real patients' data has a direct impact on the capacity of EHR systems to generate synthetic data**. Synthetic structured patient records are then generated, by using constrained statistical or machine-learning methods that preserve population-level distributions and clinically meaningful correlations while preventing any one-to-one correspondence with actual patients. Clinical constraints (such as physiological ranges, valid disease-medication combinations, and temporal coherence) must be explicitly enforced so as to avoid the creation of medically implausible records.

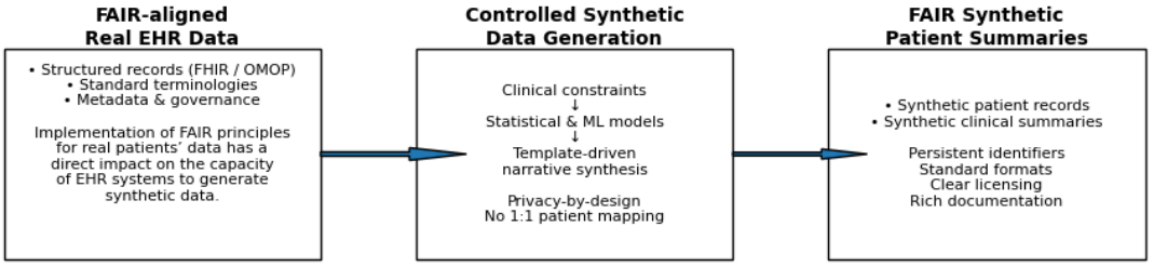


Figure 10: Relationship between FAIR implementation in EHR systems and the generation of FAIR synthetic patient summaries

Once generated, synthetic patient summaries must undergo validation at three levels:

- **Statistical validation** ensures that the synthetic population reflects the intended distributions and correlations of the source data without reproducing individual records.
- **Clinical validation** verifies that the summaries are internally consistent, medically plausible, and aligned with standard care pathways.
- **Privacy validation** confirms that the text contains no personal identifiers, rare combinations, or distinctive phrasing that could allow inference about actual individuals. Automated checks are typically complemented by expert clinical review.

To comply with the FAIR principles, the resulting dataset must be accompanied by rich metadata describing the scope of the data, the generation methodology, the validation steps, and the intended use cases. Persistent identifiers and open, standardised formats make the data findable and accessible, while the use of healthcare terminologies and standard schemas ensures interoperability. Clear licensing terms, versioning, and documentation of limitations are essential to make the synthetic patient summaries reusable for research and algorithm development, while explicitly excluding clinical decision-making or patient care.

Finally, a disclaimer must always be presented (see the following example). "This dataset consists entirely of synthetic patient summaries generated by controlled templates and language models. It contains no real patient data, cannot be traced to any individual, and is provided solely for research, development, and educational purposes. It is not suitable for clinical decision-making or any application involving real personal health information."[79]."

The HL7 SYNDERAI initiative provides a good example of how synthetic data can be used for different purposes (in this case, displaying both EU and European Member State (and other nation) Patient Summaries)[80].

12 Recommendations

The recommendations presented in this section have been adapted to consider the inputs received during the two workshops organised in 2025 (see annexes 3 and 4).

To date, 10 recommendations have been formulated which could lead to a calibrated and carefully governed approach to algorithm-based tools that are relevant to the EHDS and EHR systems.

12.1 Implement deep semantic interoperability strategies

Move beyond superficial data exchange towards **semantic interoperability "by design"** i.e., intentionally. This means to:

- Adopt international standards like FHIR® for structuring clinical resources,
- Use robust ontologies and coded vocabularies (e.g., SNOMED CT, LOINC, ICD),
- Invest in national terminology servers (as has been done in Austria, France, and the Netherlands) to ensure alignment, translation, and dynamic updating of standards.

Semantic strategies must also increasingly cover multimodal inputs (e.g., imaging, omics, signals, and unstructured text), aligning with DICOM, whole slide imagining (WSI), and national exchange formats.

[79] The example has been created manually but is inspired by examples used in several EU projects focusing on AI-CDSS.
[80] https://synderai.net/index.php?menu=examples/EPS

12.2 Prioritise data quality assurance

Establish a Three-Pillar Foundation: FAIR → Data Fluidity → Data Quality Assurance

Use principles emerging from the QUANTUM project to establish **comprehensive data quality frameworks** covering:

- Technical attributes (e.g., accuracy, completeness, timeliness, consistency, provenance, and recency),
- Operational maturity (automated processes under human stewardship),
- Fit-for-use and fit-for-purpose assessments tailored to both clinical care and algorithm readiness.

This approach will require having dedicated data governance teams, systematic quality labelling, and adaptable frameworks that evolve with local needs.

High-risk AI systems (such as AI-CDSS) must demonstrate:

- Representative, relevant, error-minimised datasets,
- Transparency on data provenance, metadata, and documentation,
- Ongoing bias monitoring and fairness monitoring,
- Continuous quality tracking as part of post-market surveillance (cf. the MDR and AI Act).

Robust data-quality assurance (including quality labels e.g., bronze/silver/gold, auditability, and dataset readiness reviews) is a priority prerequisite for safe AI deployment.

12.3 Implement a three-level Tiered Data Verification Framework

Promote Progressive – i.e., not total – Data Verification: not all data must be verified to the highest level. “Sufficient verification” needs to be defined by context and risk.

The framework must be context-specific, taking into account the level of risk, and therefore restricting high levels of constraint (e.g., strict quality, provenance, and validation) to high-risk decisions. The framework also needs to include adapted thresholds for data completeness, recency, and coding consistency, together with essential performance requirements for AI DSS.

Three levels of data verification are proposed:

- **Low risk (documentation assistance)**: focusing on structural checks, basic quality metrics, and clinician oversight.
- **Medium-risk (e.g., triage support)**: structured data completeness, provenance metadata, validation against benchmark datasets.
- **High-risk (e.g., medication decisions, diagnostics)**: which must have a full provenance chain, local dataset validation, and strict accuracy thresholds.

12.4 Implement certification of both AI systems and Deployment-Site Dataset Readiness

Certification should cover a number of areas: the AI model (safety, robustness, and explainability), local dataset (structure, mapping, completeness, and quality metrics), integration workflows, and monitoring parameters (bias drift and model performance). Several existing and emerging frameworks can support the implementation of certification for both AI systems and deployment-site dataset readiness. At the regulatory level, the EU AI Act and the Medical Device Regulation (MDR) provide a mandatory backbone covering data governance, risk management, human oversight, transparency, and post-market monitoring for high-risk medical AI. These requirements can be operationalised using established standards such as ISO 14971 for risk management, complemented by AI-specific extensions addressing bias, model drift, and robustness. For dataset readiness, and aside from the QUANTUM project outputs mentioned earlier, domain-specific frameworks such as the METRIC data quality framework[81] offer structured methods to assess dataset structure, completeness, representativeness, mapping, and quality, aligning with AI Act data-quality obligations. Trustworthy-AI guidance such as the FUTURE-AI framework[82] further supports certification by defining practical criteria for fairness, explainability, robustness, usability, and traceability across the AI lifecycle. Together, these frameworks can be combined into an integrated certification approach covering the AI model, local deployment datasets, workflow integration, and continuous monitoring of performance and bias, suitable for implementation by developers, healthcare providers, and conformity-assessment bodies in the EU context.

12.5 Integrate user-centred design and continuous co-creation

Shift from static "one-time" EHR rollouts to **iterative co-design processes** involving clinicians, nurses, and even patients and their carers – whether they are formal or informal carers.

Incorporate tools like:

- Usability logs and user feedback loops,
- Adaptive interfaces that evolve with clinical needs and roles,
- Human-in-the-loop reviews of AI outputs to maintain trust and safety,
- Stratified alerts,
- Transparency dashboards,
- Controls for turning off non-critical notifications.

[81] Schwabe D, Becker K, Seyferth M, Klaß A, Schaeffter T. The METRIC-framework for assessing data quality for trustworthy AI in medicine: a systematic review. npj Digital Medicine. 2024;7:203
[82] https://future-ai.eu/

User-centred design is also crucial, to ensure that structured data capture becomes effortless and meaningful, by being embedded seamlessly into clinical workflows.

Feedback from the 5 December 2025 workshop reinforced the point that user-centred design must address clinical alert fatigue, ensure intuitive verification loops, and make structured data capture effortless for clinicians.

12.6 Leverage AI for progressive FAIRification

FAIRification of EHR data is a stepwise approach.

AI can improve and structure the EHR data itself. Hence, once a baseline level of data structure is achieved, AI can help accelerate FAIRification. For instance:

- NLP tools can semi-automatically structure free-text clinical notes, detect inconsistencies, reduce clinician burden and summarise patient-generated data,
- Learning systems can iteratively improve data extraction by replacing human validation,
- Hybrid approaches can balance pre-coordination and post-coordination of complex concepts (e.g., by combining simpler SNOMED CT codes at runtime instead of forcing rigid up-front definitions).
- Federated learning setups that preserve data locality while enriching model quality are important facilitators.

12.7 Invest strategically at multiple levels

Full FAIRification is not achieved by an individual healthcare organisation alone. FAIRification also requires:

- National/regional investments in core infrastructures (terminology servers, validated medicinal product datasets under the ISO IDMP[83], sovereign cloud services, and semantic frameworks),
- Organisational strategies for the staffing of expert data stewards and setting up agile data quality teams,
- EHR vendors evolving to support open, modular integration and external algorithmic tools with API conformance testing,
- The promulgation of anti-lock-in rules and procurement guidelines requiring interoperability,
- The definition of National Governance Model for Data Quality and Semantics including roles (e.g., national semantic hub, hospitals, vendors, AI developers), responsibilities

[83] https://www.ema.europa.eu/en/human-regulatory-overview/research-development/data-medicines-iso-idmp-standards-overview

(mapping, quality monitoring, and audit), update cycles, and accountability mechanisms.

12.8 Balance transparency, explainability, and seamless AI

There are several paradoxes to resolve:

- To be **trusted and adopted**, AI must be explainable and keep clinicians "in the loop".
- To ensure **daily operational acceptance**, however, AI must "disappear into the background" ("be invisible and seamless"), having removed administrative burdens without providing distracting details.

Designing for both explainability and transparency ("invisibility") will be key. To achieve this, high-risk AI systems must provide:

- Clear explanations of logic (appropriate for clinicians),
- Communication of uncertainty,
- Full traceability of inputs/outputs,
- User awareness when interacting with AI (a transparency obligation).

EHR vendors must support:

- Open integration with third-party AI tools via standardised APIs,
- Model versioning, traceability, and rollback capabilities,
- Privacy-preserving deployment setups (often preferring EU-sovereign clouds),
- Change management processes consistent with MDR "significant change" rules.

12.9 Recognise that advanced AI algorithm integration is still nascent

While some advanced health systems (examples which originated largely from the USA include Mount Sinai, Kaiser Permanente, Mayo Clinic, and EPIC and Cerner clients) have begun embedding AI into their systems, they are exceptions and are often aided by having their own integrated ecosystems. In contrast, many European settings still face foundational hurdles with establishing interoperability and data quality.

AI-CDSS must support techniques such as drift detection, incident reporting, and periodic performance reviews which are aligned with the MDR and AI Act post-market surveillance requirements.

Thus, a phased approach is essential – starting by improving fundamental data structures and quality before adding additional layers of complex AI tools.

12.10 Invest in adapted stakeholder specific digital literacy

Develop "FAIR" factsheets, which are adapted to each type of stakeholder category, and which support clinicians with structured coding training i.e., tools that make structured capture usable. Develop strategies to improve the understanding of data quality metrics and data

literacy so that e.g., clinicians can interpret accurately any recommendations made by AI tools.

13 Conclusions

The European Health Data Space (EHDS) is expected to make a significant contribution to the FAIRification of health data, notably through the interoperability specifications defined for the six priority information domains listed in Article 14 of the Regulation. By promoting common formats, identifiers, and exchange mechanisms, the EHDS establishes an essential foundation for making health data more findable, accessible, interoperable, and reusable across borders and systems.

However, these specifications alone may be insufficient to support the safe, scalable, and trustworthy integration of more advanced algorithm-based tools, particularly AI-driven clinical decision support systems (AI-CDSS). While syntactic interoperability enables data exchange, the effective reuse of data by algorithms requires a deeper and more systematic implementation of FAIR principles, especially semantic interoperability, data provenance, contextual metadata, and versioning. Without these elements, data remain difficult to interpret consistently, reuse across contexts, or validate for high-risk clinical applications.

A deeper implementation of semantic interoperability is therefore a critical prerequisite for meaningful communication between EHR systems and algorithmic tools. This activity includes a consistent use of shared clinical information models, standard terminologies, traceable mappings, and machine-readable metadata describing data quality, fitness-for-purpose, and usage constraints.

Achieving this level of FAIRness in Europe requires coordinated strategies across the entire health data value chain – at national, regional, and local levels – and sustained investment in both specialised human resources (such as data stewards and semantic experts) and in enabling technical infrastructures.

Artificial intelligence (AI) is increasingly becoming part of the solution to FAIRification challenges. AI-based tools can support the structuring, coding, validation, and enrichment of health data, thereby reducing reliance on manual processes and mitigating clinician burden. At the same time, the deployment of such tools must preserve human oversight, ensuring that clinicians remain “in the loop” for ethical, legal, and clinical quality reasons. From a FAIR perspective, this human oversight is also essential to maintain transparency, explainability, and trust in both data and algorithmic outputs.

The EHR industry continues to play a central role as a gatekeeper in the integration of AI tools, shaping the degree to which data can be accessed, reused, and combined across systems. Ensuring that EHR platforms support open interfaces, standardised APIs, and

traceable data flows is therefore essential to avoid lock-in effects and to achieve completely the FAIR objectives underpinning the EHDS.

A number of algorithm-based tools require only loose integration with EHR systems and can already deliver value by relying on well-defined, FAIR-enough data subsets. These tools may be easier to deploy in the short term, as they demand minimal interaction from end-users and can help reduce administrative workload. However, their long-term scalability, safety, and cross-context reuse will still depend on the progressive FAIRification of underlying data structures. As such, incremental deployment strategies should be aligned alongside a clear roadmap toward more comprehensive FAIR data maturity.

Looking ahead, the progressive implementation of FAIR data principles will be a decisive enabler for compliance with the EU Artificial Intelligence Act (AI Act) and the Medical Device Regulation (MDR). Both regulatory frameworks place strong emphasis on data quality, representativeness, traceability, transparency, and post-market monitoring – requirements that cannot be met sustainably without FAIR-aligned data structures. High-risk AI systems, including AI-based clinical decision support tools regulated under the AI Act and MDR, depend on datasets whose provenance, semantic meaning, quality metrics, and contextual limitations are explicitly documented and machine-readable. In this sense, to conclude, FAIR data should not be viewed solely as a technical or interoperability objective, but as a regulatory and safety prerequisite.

Embedding FAIR principles into EHR systems and governance frameworks will be essential to enable trustworthy AI deployment, facilitate conformity assessments, support post-market surveillance, and ensure that innovation in digital health remains aligned with European values of safety, accountability, and human oversight.

14 Annex 1:
Typology of algorithm-based tools for EHR systems

The main algorithmic tools that can interact with EHR systems span several fields, including data analytics, artificial intelligence, and process automation. They enable enhanced clinical decision-making, predictive analytics, and personalised care. These tools use advanced algorithms such as machine learning, natural language processing (NLP), and data analytics to extract value from EHR data.

When integrated in EHRs, these algorithmic tools can improve the quality of care, increase operational efficiency, and provide valuable information to healthcare professionals.

We list below the 10 most commonly used types of algorithms and their uses. The order in the list is random

- **Machine Learning Algorithms:**
  - **Supervised Learning**: To predict clinical outcomes, diagnose diseases, or recommend treatments.
  - **Unsupervised Learning:** To identify trends or groups of similar patients in the data being learned on.
- **Natural Language Processing (NLP):**
  - **Information Extraction:** To extract structured data from unstructured clinical notes.
  - **Sentiment Analysis:** To understand patients affective states based on their comments or ratings.
- **Computer Vision Algorithms:**
  - **Medical Image Analysis:** To interpret X-Rays, Magnetic Resonance Imaging (MRI)s, or other medical images.
- **Recommendation Algorithms (also known as Recommender Systems):**
  - **Treatment Recommendations:** To suggest treatments based on historic patient data.
  - **Drug Recommendations:** To suggest medications based on medical history and drug interactions.

- **Anomaly Detection Algorithms:**
  - **Patient Monitoring:** To detect abnormalities in vital signs or other health data in real time.
- **Prediction Algorithms:**
  - **Risk Prediction:** To estimate the risk of chronic diseases, hospital readmissions, or other clinical events.
- **Optimisation Algorithms:**
  - **Resource Planning:** To optimise the use of hospital resources, such as beds or operating rooms.
  - **Inventory Management:** To manage inventory of medications and medical equipment.
- **Security and Privacy Algorithms:**
  - **Data Encryption:** To protect sensitive patient information.
  - **Intrusion Detection**: To monitor and prevent unauthorised access to patient records.
- **Data Mining Algorithms:**
  - **Data Exploration:** To discover patterns and relationships in large healthcare databases.
- **Data Fusion Algorithms:**
  - **Data Integration:** To combine data from different sources (e.g., medical records, wearables, genomic databases).

15 Annex 2: A sample of tools deployed in EHR systems

This annex contains a list and description of the key algorithm-based tools that can be integrated with EHR systems. For each category, we provide a description of their general purpose, and a number of examples of products already available on the market.

The availability of algorithmic-based tools is, of course, quite a new and extremely dynamic area. As a consequence, there is difficulty in identifying which products which are "here to stay".

This list is therefore not exhaustive, and the categories proposed are suggested purely as examples. Tools originate from and are used in the USA, Canada, or in different nations in Europe. In several cases, the use cited for the tool takes place simply in a single location. In other cases, recent or ongoing criticisms of the tool(s) are included in the detail provided.

1. Clinical Decision Support Systems (CDSS)

- **Purpose**: To assist healthcare providers in making clinical decisions by providing evidence-based recommendations and alerts based on patient data.
- **Examples**:
  - **Cerner PowerChart**: Is integrated with EHRs to deliver clinical decision support related to medication dosages, clinical guidelines, and potential drug interactions.
  - **Duodecim/ Evidence linker**: Scripts are integrated with general practices' EHRs and the linker provides recommendations based on ICPC2[84] codes. The scripts are however currently more rules-based than machine learning-based. The scripts originally developed in Finland have been used and implemented in other countries, e.g., in Belgium.
    - https://confluence.ihtsdotools.org/display/doccds/ebmpracticenet
    - https://www.evidencelinker.be/fr
  - **UpToDate**: Provides evidence-based clinical decision support, using algorithms to suggest treatment guidelines based on patient conditions; its value is directly tied to the accessibility and interoperability of clinical data. UptoDate is today one of the most widely used tools, in particular by healthcare specialists who are active in research.

[84] https://www.who.int/standards/classifications/other-classifications/international-classification-of-primary-care

- **IBM Watson for Oncology**: Leverages AI to analyse clinical data and assist in oncology decision-making. FAIR data ensures that Watson can access and interpret structured clinical data, such as diagnostic codes, medications, and lab results, in a standardised way. The system was, however, abandoned in 2023, due to the suspicion that Watson provided inappropriate or even unsafe recommendations.

2. Predictive Analytics Tools

- **Purpose**: To use machine learning algorithms to predict patient outcomes, such as the likelihood of readmission, disease progression or risk of adverse events. These predictions depend on clean, standardised, and accessible historical data from EHRs to identify trends, correlations, and outliers.
- **Examples**:
  - **Health Catalyst**: Uses predictive analytics and machine learning algorithms to identify patients at risk of various conditions, helping with proactive care management.
  - **Mede Analytics**: Offers predictive analytics solutions, helping providers predict hospital readmissions, patient deterioration, and other clinical events. Access to interoperable and reusable data from different sources ensures the tool can deliver accurate predictions.
  - **Qventus**: A real-time hospital operations AI platform that uses predictive analytics to improve patient flow, reduce readmissions, and manage hospital resources. FAIR data is essential for accurate predictions because data needs to be findable, accessible, and of high quality to train predictive models.

3. Natural Language Processing (NLP) Tools

- **Purpose**: To analyse unstructured data (e.g., physician notes, discharge summaries, radiology reports) in EHRs using NLP techniques to extract relevant information for clinical decision-making. For NLP to function effectively, the unstructured data needs to be accessible, well-annotated, and standardised (at least at metadata level) to enable accurate processing and analysis.
- **Examples**:
  - **Clinical Text Analysis and Knowledge Extraction System (Apache cTAKES)**: An open-source NLP tool designed for extracting information from clinical text within EHR systems.

- DeepMind Health **(Google Health)**: Uses NLP to analyse EHR data for predicting disease outcomes and detecting early signs of illnesses such as kidney disease.
- Nuance's Dragon Medical One: A speech-to-text platform that uses NLP to extract structured clinical data from unstructured physician notes.

4. Medical Imaging Analysis Tools

- **Purpose**: To integrate EHR data with medical imaging data, using AI and machine learning algorithms to analyse radiological images, pathology slides, and other diagnostic images for disease detection and progression tracking. These tools need to access high-quality imaging data, which must be structured and standardised to ensure compatibility with algorithmic analysis. This segment of tools has been the one which has registered the highest rate of adoption by clinical organisations to date.
- **Examples**:
  - Aidoc: An AI-powered radiology tool that assists in detecting acute conditions such as haemorrhages, strokes, and pulmonary embolisms, integrated with EHRs for seamless analysis and alerts. Aidoc claims to have an important comparative advantage as the platform provides a unified user interface for accessing results from both Aidoc and third-party algorithms, thereby promoting a better experience for users.
  - Second opinions**:** This platform collaborates with Nanox AI **(**which helps to identify patients who are asymptomatic or who have undetected symptoms correlated with chronic cardiac, liver and bone conditions) in promoting preventive care management. It uses machine learning algorithms to analyse medical imaging data, such as X-rays and computed tomography (CT) scans, and provides diagnostic insights integrated with EHR systems.
  - Viz.ai: An AI tool that analyses CT scans for signs of stroke and is integrated with an EHR to notify physicians in real-time.

5. Clinical Pathway Tools

- **Purpose**: To guide clinicians through a series of evidence-based clinical practices for managing specific medical conditions, ensuring standardised treatment pathways.
- **Examples**:
  - Carealign: Helps clinicians implement evidence-based clinical pathways that are integrated into the EHR workflow.

- **Pathfinder by Elsevier**: Uses algorithms to support clinicians in adhering to best practices and clinical pathways for treating specific diseases.

6. Drug-Drug Interaction (DDI) Checkers and Medication Safety Tools

- **Purpose**: To analyse medication prescriptions in real-time to identify potentially harmful drug interactions, dosages, or adverse effects based on patient data in the EHR. For accurate alerts and recommendations, data about medications, dosages, allergies, and patient conditions must be accessible, interoperable, and standardised.
- **Examples**:
  - **Drugs.com DDI Checker**: Provides real-time alerts about potential drug interactions by analysing the patient's medication list in the EHR.
  - **Micromedex**: Offers drug interaction alerts and evidence-based clinical information, which integrates with EHR systems to ensure safe medication prescriptions.
  - **Synbase**: Enables the integration of drug-drug interactions into electronic health record or pharmacy information system.

7. Population Health Management Tools

- **Purpose**: To use data from EHR systems to monitor and manage the health of entire populations by identifying trends, gaps in care, and patients who need intervention. For accurate insights and interventions, the data must be interoperable across systems and accessible in standardised formats.
- **Examples**:
  - **Arcadia**: An integrated platform that helps healthcare organisations manage population health using data-driven insights from EHR systems to improve outcomes and reduce costs.
  - **Emed**: eMed Empathetic AI™ is the leader in Population Health GLP-1[85]/GIP[86] programmes and chronic care management related to obesity and type 2 diabetes.
  - **HealthEC**: A population health management platform that uses predictive analytics and data from EHRs to identify high-risk patients and coordinate care.

[85] GLP-1 is a type of medication to be taken when a person has type 2 diabetes: https://www.diabetes.org.uk/about-diabetes/looking-after-diabetes/treatments/tablets-and-medication/glp-1.
[86] A gastric inhibitory peptide (GIP) is a hormone that stimulates the release of insulin on the oral intake of food.

8. Frailty and Risk Stratification Tools

- **Purpose**: To predict which patients are at higher risk of adverse health outcomes, such as falls, complications, or hospitalisation, by analysing various EHR data points. These tools assess the frailty, or overall health risk, of patients, often by using EHR data points like age, comorbidities, and laboratory results. Standardised and accessible data ensures that these tools can make accurate predictions about a patient's risk level.
- **Examples**:
    - **Elder Care Analytics (ElderCheck)**: A tool designed for analysing risk factors in older adults and assisting in predicting and preventing falls, cognitive decline, and hospitalisation. To function effectively, it needs standardised, accessible data from EHRs.
    - **Geriatric Risk Stratification Tools**: These tools (e.g., **Frailty Index**[87]) integrate with EHRs to assess the frailty of elderly patients and help clinicians make better care decisions. The value of this tool is heavily dependent on the quality and interoperability of the data.

9. Chronic Disease Management Tools

- **Purpose**: To help in monitoring and managing chronic conditions such as diabetes, hypertension, and cardiovascular diseases through real-time data collection, analysis, and intervention planning.
- **Examples**:
    - My Diabetes Home: A diabetes management tool that integrates with EHRs to track blood sugar, insulin use, and other relevant data, offering personalised insights.
    - Omron Connect: Integrates with EHRs and uses algorithms to track patient data related to hypertension, offering suggestions for managing the condition. FAIR data from an EHR, such as blood pressure readings, is essential for accurate monitoring and intervention recommendations.

10. Virtual Health Assistants

- **Purpose**: To use AI algorithms to offer personalised healthcare recommendations, reminders, and information to patients, while integrating with EHRs to update patient records and track progress. To offer such personalised insights, an assistant requires real-time access to standardised EHR data. The core capabilities of AI EHR Assistants

[87] E.g., https://www.england.nhs.uk/ourwork/clinical-policy/older-people/frailty/efi/.

are voice recognition, contextual understanding, real-time clinical insights, and customisable workflows.

- **Examples**:
    - **Ada Health**: An AI-driven health assessment tool that integrates with EHRs to provide symptom checking and personalised health insights. This assistant has been developed first of all for the patients, but it is integrated with EHRs by making patient complaints and history in clinical terminology available to physicians and includes them in their workflow.
    - **Suki Assistant**: Suki provides ambient documentation, dictation, ICD-10[88] and hierarchical condition category (HCC) coding, and answers questions. It has bidirectional, read/write capabilities with all leading EHRs – it operates mainly in the USA.

[88] International classification of diseases (ICD) 10 https://icd.who.int/browse10/2019/en.

16 Annex 3: Report of the first workshop (October 14, 2025)

This annex outlines the main aspects of the October 14, 2025 workshop and its chief concerns and take-aways.

16.1 Meeting context and objectives

This workshop took place in the context of the EHDS Implementers Task Force created by EHTEL to create awareness and stimulate discussions among EHTEL members and contacts about the implementation of the EHDS regulation.

The meeting took place on 14 October 2025 and was moderated by Heidi Peltotalo (THL-Finland) and Daniela Spießberger (Gematik-Germany). Its main objective was to collect inputs from EHR and AI decision support systems developers and have a first feed-back on the recommendations and questions listed in a short briefing paper which was shared with the audience.

The three speakers were Luc Nicolas (EHTEL), Dragan Sahpaski (Sorsix), and Stefano Tagliaferri (Keylon).

The format included an introductory presentation of the briefing paper, the two developers' interventions, an interactive poll with participants, and an open discussion.

16.2 Participation

This first workshop was designed for a limited and selected audience.

A total of 38 people (excluding organisers) from 13 different countries registered for the workshop with a balanced representation of public, private, and non-profit sectors.

	Non-profit	Private	Public	Total
Belgium	1			1
Czech Republic			1	1
Denmark	1		1	2
Finland			2	2
Germany	2	2		4
Greece	1			1
Hungary			1	1
Italy	5	7	3	15
Luxembourg			1	1

	Non-profit	Private	Public	Total
North Macedonia		2	1	3
Norway	1			1
Scotland			2	2
Spain			2	2
United Kingdom	1		1	2
Total	**12**	**11**	**15**	**38**

Table 2: Participation statistics - workshop 14 October 2025

Twenty-one people actually attended the workshop. The highest no-show rate was for both public and non-profit sectors averaging 50%, while the private sector showed a 70% participation rate. Representatives from seven countries eventually attended the meeting.

	No-show	Participants	Total
Belgium	1		1
Czech Republic	1		1
Denmark		2	2
Finland		2	2
Germany		4	4
Greece	1		1
Hungary		**1**	**1**
Italy	**7**	**8**	**15**
Luxembourg	1		1
North Macedonia	2	1	3
Norway	1		1
Scotland		2	2
Spain	2		2
UK	1	1	2
Total	**17**	**21**	**38**
	44.74%	55.26%	

Table 3: Participation statistics - show/no-show - workshop 14 October 2025

16.3 Introductory presentation

Luc Nicolas (EHTEL) provided an introduction of the briefing paper and emphasised specifically the following aspects:

- **AI is already a reality in daily life:** The healthcare challenge is not *if* AI will be used but *how* to keep pace with its evolution.
- **The concept of data fluidity is a key enabler:** Health data should be created and structured with the intention to flow for multiple purposes – from individual care to societal benefit. It implies both interoperability and generic reusability.
- **Data quality assurance is a must:** Current structured/coded data are often produced for administrative or reimbursement purposes rather than clinical care and hence are considered as trusted data by clinicians. Clinician trust and semantic quality need to be strengthened at the point of care.
- **AI as opportunity or bottleneck:** Innovative SMEs are driving development, but entrenched EHR vendors can also act as gatekeepers.

Luc Nicolas then introduced the main barriers identified and commented shortly on the proposed recommendations and associated questions.

16.4 Developers' presentations

- **Dragan Sahpaski (Sorsix)** is the chief technology officer (CTO) of Sorsix, a company headquartered in Australia with a development centre in North Macedonia. The EHR "Pinga" platform is deployed in 10+ countries – it is a modular SaaS EHR/HIS with over three billion clinical interactions per year. Sorsix is active in EHDS interoperability testing (such as at the annual Integrating the Healthcare Enterprise (IHE) Connectathon[89]).
    - **Dragan emphasised that AI has to be seen both as a data consumer and catalyst**: This ranges from NLP structuring, automated coding, data harmonisation to predictive analytics, clinical decision support and operational optimisation.
    - The major **challenges** he identified were:
        - **Data quality gaps:** Poor processes undermine model reliability.
        - **Semantic interoperability:** Huge variability in clinician coding (e.g,. ICD, SNOMED CT, local codes).
        - **Vendor dynamics:** Legacy vendors can block innovation; even if FHIR® is increasingly adopted, its real adoption often remains slow.
        - **User-centric design:** Most legacy EHRs are administrative, not clinical.

[89] https://www.ihe.net/testing/connectathon/

  - His **main messages** were:
    - EHDS is both an opportunity and a **systemic transformation challenge**.
    - AI integration is only as strong as underlying **data quality**.
    - Vendors must act as **bridges, not gatekeepers**.
    - Agile **change management**, rapid data model evolution, and clinician engagement are needed.
    - Regulation should **set a minimum bar**, not dictate innovation pathways.

- **Stefano Tagliaferri** is the CTO of Keylon, a company based in Italy, which develops several modular AI-supported decision support systems meant to interact with EHR systems.
  - **Stefano emphasised the need for a pro-active collaboration between AI and EHR vendors** which is not only necessary but also beneficial to both sets of developers. For this collaboration to take place, he insisted on frictionless integration though the use of plug-and-play- standards-based (e.g., HL7, DICOM) APIs. Adapted business models need also to be developed (e.g., co-branding, revenue sharing, mutual value creation). He mentioned as the main integration barrier the fragmentation of national systems and legacy interfaces.
  - His **main messages** were:
    - Everything starts with **solid data governance**.
    - Implement **pilot projects** in high-impact areas (e.g., radiology, triage) where data are also usually quite well structured.
    - Engage **end-users** early so as to avoid resistance and demonstrate with them the benefits (time saved, fewer errors, better outcomes).
    - Be logical with your procurement policy and select only **open, interoperable platforms.**
    - AI can simultaneously help with the structuring data and enable decision support, reducing clinician workload while improving care quality. First examples of progress are expected in clinical documentation and data structuring (voice recognition, NLP), clinical summarisation tools, and segments where data are already well structured. Complex diagnostic/therapeutic support is delayed by regulatory processes (e.g., the AI Act, MDR/In Vitro Diagnostic Regulation (IVDR) certifications) but also due to the need of data validation and clinical acceptance. Europe's slower adoption pace is due to strict certification timelines.

16.5 Interactive poll and key discussion points

Participants voted on 12 key questions (relevance was rated 1-5: here this has been translated into the expressions “very high” or “high” to “mixed”. Any item lower than mixed is not report here):

Theme	Questions assessed	Average relevance rating
Certification	Should the certification of AI-based tools include assessments of the quality and structure of local datasets where they will be deployed?	Very high
Cross-border	How do we address the risks of AI tools misinterpreting data when they are moved across institutions or countries with different data practices?	Very high
Data Quality	Given the “garbage in, garbage out” dilemma, how can we ensure that the data used by AI tools is of sufficient quality and context?	High
AI in for AI output	Is the deployment of AI decision support tools conditioned by the use of AI data FAIRification tools?	High
Governance	What governance models can support continuous data quality improvement and iterative algorithm design in clinical settings?	High
Co-creation	Why does co-creation with clinicians often fail in practice despite being part of project rhetoric?	High
EHR vendors	Are current EHR vendors enablers or blockers of innovation in AI-driven healthcare? How can this dynamic be modified?	High
Early Deployment	Which are the AI tools which are meant to be deployed the quickest? To support which type of process? By whom and where?	High
Explainability	Can AI be both explainable and ‘invisible’ in clinical workflows? Or are these goals inherently contradictory?	High
Regulation	Should there be regulatory mechanisms to compel EHR vendors to open up interfaces for third-party AI tools?	Mixed
Early Adopters	What can we learn from early adopters like the Mayo Clinic? And how transferable are their models to European contexts?	Mixed
Semantic Standards	Do current semantic standards (e.g., SNOMED CT, LOINC, ICD) sufficiently meet the needs of AI integration. Or do we need new approaches?	Mixed

Table 4: Assessment of questions during 14 October 2025 workshop

The key concerns which surfaced from the discussion were:

- Semantic interoperability is **still weak**.
- Trust in data and governance **is central**.
- Regulation versus innovation balance is **a sensitive area**. One should ensure regulatory frameworks encourage innovation rather than act solely as constraints.
- Clinician involvement is **critical for successful AI adoption**.
- Integrated ecosystems like those of the Mayo Clinic and some Israeli HMOs enable faster innovation, but it is unclear how transferable their models are to EU systems, which are more fragmented.

16.6 Key take-aways

- **AI and EHR integration are inevitable** – but success depends on data quality, governance, and openness.
- **Semantic interoperability remains a critical bottleneck**.
- **Regulation alone is insufficient**; incentives and collaboration with vendors are key.
- **Clinician engagement** is necessary to ensure practical adoption.
- The **first AI solutions** to take off will likely be pragmatic – documentation and workflow tools – rather than complex diagnostics.

17 Annex 4: Report of the second workshop (December 5, 2025)

This annex records the chief aspects of the December 5, 2025 workshop, and those aspects of it which required further development (i.e., to be transformed into recommendations), and may still need ongoing development.

17.1 Workshop context and objectives

This second workshop was held on December 5, 2025. It involved a different and larger community than for the first one because, going beyond EHTEL membership and the EHTEL network, invitation to participate had been shared with all Xt-EHR partners and other European projects which have a direct relationship with EHDS implementation, namely xShare, XIA, I2X and MyHealth@Myhands.

The workshop's main objective was to revalidate and/or add nuance to the key recommendations of the briefing paper with a larger engaged community and discuss the FAIR requirements more in detail through support of two previously prepared use cases.

The workshop was introduced by Zoltan Lantos (representing Xt-EHR) and moderated by Dragan Sahpaski (Sorsix) and Luc Nicolas (EHTEL).

The format of the workshop included an updated presentation of the short briefing paper, followed by two break-out sessions focusing on the use cases, and ended with a plenary to collect key messages.

17.2 Participation

A total of 166 people registered for the workshop. Invitations to attend were shared both through projects' private channels and EHTEL public channels. While attendees may have been associated with membership of multiple projects, only their main association is used here for statistical purposes.

A little more than half of all people who registered were affiliated with the Xt-EHR Joint Action. Roughly one-third of the people registered eventually did not attend the workshop. People from the xShare project had a much higher level of no-show.

Project	Registration	Attendance	%	No-show
Xt-EHR	86	53	51.81%	38.37%
xShare	23	9	13.86%	60.87%
MyHealth@MyHands	12	7	7.23%	41.67%
Other	11	10	6.63%	9.09%
XIA	10	7	6.02%	30.00%
i2X	9	6	5.42%	33.33%
EHTEL members	9	8	5.42%	11.11%
Organisers	6	6	3.61%	0.00%
Total	**166**	**106**	**63.85%**	**36.15%**

Table 5: Participation statistics (projects) - 5 December 2025 workshop

People from 27 countries registered: Germany, Greece, Hungary, Italy, Portugal and Spain, each had 10 or more registrations. Croatia, Slovenia, and Slovakia had only one person registered, but those who had registered did not actually attend the meeting. Thus, ultimately, people from only 24 countries actually participated. The level of no-show was spread roughly across countries, with only Austria, France, Czech Republic, and Latvia recording a no-show rate above 50%.

Looking at the type of stakeholders, 48% were affiliated to public entities while 28% were associated with the non-profit sector (research, healthcare organisations) and 24% with the private sector. The level of no-show was higher for public entities (41%) than for the non-profit sector (34%) and private actors (31%).

Country	Non-profit	Private	Public	Total
Austria	1		1	2
Belgium	6	4	5	15
Croatia			1	1
Cyprus	5	1	3	9

Country	Non-profit	Private	Public	Total
Czech Republic		1	3	4
Denmark	2		2	4
Estonia			1	1
Finland			8	8
France	1	2	3	6
Germany	6	4	1	11
Greece	2	6	2	10
Hungary	1	3	7	11
Italy	7	6	7	20
Latvia			2	2
Netherlands			2	2
North Macedonia	1	1		2
Norway	1	2	3	6
Poland	1			1
Portugal	3	4	6	13
Republic of Ireland			7	7
Romania			3	3
Slovakia			1	1
Slovenia	1		1	2
Spain	8	3	8	19
Sweden			3	3
United Kingdom	1	1		2
United States		1		1
Total	**47**	**39**	**80**	**166**

Table 6: Participation statistics (countries) - 5 December 2025 workshop

17.3 Presentation of the Working Paper updated positions

The working paper was shared with the participants ahead of the meeting. It provided a brief but consolidated analysis of how EHR systems should evolve to support high-quality, safe, interoperable, and AI-enabled clinical workflows in the EHDS.

The paper's primary pillars included FAIR data, data fluidity, interoperability, ethical AI deployment, and the role of vendors. By December 2025, the October version of the paper had been updated with considerations about three deployment options: **On-Premises AI** (local computation, sensitive data protected); **Cloud AI** (high compute capacity, cross-border capabilities); **Hybrid Models** (pre-training in cloud + local inference). The paper does not prioritise any one of these three models, presenting them instead as equivalent choices depending on context.

The main sections of the paper (key concepts, barriers, and recommendations) were summarised by Dragan Sahpaski (Sorsix).

During the workshop, participants strongly emphasised data quality, clinical usability, vendor behaviour, and governance as areas requiring reinforced attention in the paper. They also introduced new concepts such as **tiered verification**, **dataset readiness assessment**, and **certification covering both AI systems and the datasets they run on**.

The final section of the workshop report now provides **refined, consensus-based recommendations** that merge the working paper's strategic orientation with the practical concerns expressed by workshop participants.

17.4 Clinical scenarios presented during breakout sessions

As the use cases eventually selected are already presented in detail previously in the main body of this report (see section 8 (8.1 and 8.2)), we reproduce here only some basic information.

17.4.1 Scenario 1: Type-2 Diabetes Primary Care Workflow

Persona: Maria, 62, digitally literate, uses a PHR app and a connected insulin device.
Workflow described:

- Continuous data injection from devices and apps
- AI flags moderate risk of deterioration
- GP reviews AI suggestions (dose adjustment, coaching module, scheduling tele-review)
- GP can accept, modify, dismiss, or request evidence
- Feedback loop to patient and algorithm for continuous learning.

FAIR-related requirements identified:

- Device signals linked to patient identity, standardised metadata
- Secure health data hubs
- Structured clinical fields, standardised units (e.g., glucose values), machine-readable formats
- Contextual notes (fasting status, stress), metadata licensing, longitudinal monitoring.

17.4.2 Scenario 2: Hip Replacement Pathway

Persona: Amy, 59, chronic osteoarthritis, escalating pain.
Workflow described:

- GP referral with structured data (SNOMED codes, Oxford Hip Score, X-rays, meds)
- Specialist triage and AI-assisted prioritisation
- Clinical examination, imaging, diagnosis confirmation
- Surgery booking
- AI support: structured data capture, prioritization, patient-reported outcomes tracking.

FAIR-related requirements identified:

- Structured and discoverable referral data
- Standardised metadata (imaging, scores, symptoms)
- Consistent identifiers across GPs, hospital, pharmacy
- Machine-readable labs (not PDFs).

17.5 Workshop participant feedback and suggestions

This section presents **only** participant feedback, separate from working paper content.

There is a wide consensus that the recommendations listed in the paper are **highly relevant**, namely:

- FAIR principles are essential.
- Interoperability is foundational for EU-wide AI and data reuse.
- Ethical and responsible AI deployment is necessary.
- Hybrid AI architectures represent promising pathways.
- Clinician oversight and explainability cannot be bypassed.
- Data availability remains the first barrier to AI.

The following **aspects should be more developed**:

- Clearer foundational definitions should be provided.
- More emphasis should be placed on data quality and safety issues
- Lock-in strategies should be described in more detail, highlighting structural risks.
- The paper should mention specifically risk-tiered thresholds and essential performance levels
- The governance model needed should be more detailed in calling for detailed structures, roles, and accountability
- The paper should emphasise a better articulation of regulatory layers ((liability, the Digital Omnibus (i.e., simpler digital rules)[90], oversight)
- Aside from technical aspects, the paper should pay more attention to the clinician burden, focusing on usability, workflow integration, and alert fatigue.
- More attention should be paid to contextual data (including also patient-generated data) and not only raw data streams.
- Structured awareness and training adapted to each stakeholder category should be developed.
- One should also pay attention to the importance of human oversight and patient involvement.

[90] https://commission.europa.eu/news-and-media/news/simpler-digital-rules-help-eu-businesses-grow-2025-11-19_en

Several issues mentioned (such as those related to human oversight or ethical aspects), although highly relevant, lie outside of the scope of the current working paper/report, but could be developed in a future paper.

Thus, we reproduce here comments directly related to the specific topics covered in the workshop, which have been incorporated in this report either to update or to add nuance to recommendations outlined in section 12.

- Participants repeatedly emphasised that data quality is not simply operational – it is a safety prerequisite for AI-based clinical decision support. This was among the strongest collective messages of the workshop: **poor data lead to unsafe clinical AI**. This threat to safety could be mediated via the following processes and channels:
    - Data quality must be measured systematically before deployment with a Local Data Readiness Assessment.
    - Dataset readiness should be assessed alongside AI model certification.
    - Continuous data monitoring equipped with quality labels could be deployed.
    - Provenance and context data are essential for clinical interpretation and should be included by default.
    - "Sufficient" rather than perfect verification should be the aim.
    - Patients must remain able to challenge incorrect information, which today is often difficult. Patients' validation of data also forms part of data quality assurance.
- Participants noted that AI certification is insufficient without evaluation of the data it runs on: hence, deployment-site testing should be mandatory, and monitoring should continue throughout AI use. Data certification must cover all aspects: cybersecurity, data privacy, workflow integration, explainability, bias mitigation and robustness while fit to purpose data readiness levels should be defined.
- Participants also emphasised that, today, EHR vendor lock-in is a major structural barrier (they had the following difficulties in mind): restrictions of access to structured data, limitations to API openness, delaying adoption of standards, the charging of excessive integration fees and/or creation of proprietary modules that "trap" health organisations. Participants argued that vendors are not neutral enablers and must be governed through certification, regulation, and open standards enforcement.
- Some participants also argued that traditional interoperability – standardising formats, mapping terminologies, and adding semantics after data capture through ETL – has been needed for legacy systems, but is inherently costly, "lossy" (i.e., loses data), and delivers limited value, as shown by the enormous effort required to build regional and national EHRs in many EU countries. The EHDS is seen as a unique regulatory opportunity to break with this model by mandating semantic interoperability by design:

embedding common clinical semantics at the point of data creation through a shared clinical language for healthcare. Rather than focusing on making heterogeneous EHR data interoperable after the fact, the priority should be on ensuring that vendors capture semantically consistent data from the start. This goes beyond adopting FHIR® and SNOMED CT as the action requires EHR vendors to implement shared clinical information models (such as openEHR's two-level modelling approach), and suggests the need to strengthen EHDS recommendations to mandate compliance with these models (thus, benefitting vendors, healthcare organisations, and AI) through the provision of clearer requirements, lower integration costs, and reliable semantic foundations).

- Participants proposed moving from a binary "verified versus not verified" viewpoint to context-specific verification which would take into account limiting high levels of risk/constraints (strict quality, provenance, and validation) to high-risk decisions. They also proposed to develop thresholds for data completeness, recency (the fact of being recent), and coding consistency together with essential performance requirements for AI Decision Support Systems. AI recommendations must also be traceable back to source data, model version and decision rules which imply supplementary requirements for EHR systems.
- Participants also noted that AI should reduce, not increase, documentation burden. Doctors must not be overwhelmed by notifications and hence there is a need for smart stratification, balanced alerting, and clear clinical relevance. Excessive notifications, poor AI integration into workflows, lack of explainability, unusable structured coding workflows, and the overwhelming of clinicians with patient-generated data, are important barriers that should also lead to specific requirements.
- Participants re-emphasised that AI could assist in structuring, validating, and enriching data, including e.g., autocomplete coding, correction of inconsistencies, context recognition, summarisation of patient-generated data, and intelligent prompts[91].
    - Participants mentioned the need for stronger governance and National Semantic Services; they commented that individual hospitals alone cannot perform mapping, maintain terminologies, and ensure consistent ontology updates.
- The workshop emphasised the need of bridging EHDS primary/secondary data use and AI training loops, at least for priority data categories.

[91] This topic has been covered by another EHTEL working paper, which was still under development at the time of the workshop. It has been included in Section 4 of this report.

- The use of existing datasets foreseen by Healthdata@eu should be mentioned (i.e., EHDS infrastructure for Data access and catalogues).
- The possible use and need for synthetic and pseudonymised data could be described in more detail.
- Adapted governance should define roles, responsibilities, quality ownership, ethical review, and vendor obligations.

18 Annex 5: Video-recordings of the two workshops

- Video-record of the workshop of 14 October 2025 (on YouTube, but unlisted)
- Video-record of the Workshop of 5 December 2025 (on YouTube, unlisted)