

# D6.1 Legal and ethics compliance report

## Lead Author: TLX

With contributions from: HL7, INRIA, VHIR

Reviewer: Luc Chatty (HL7), Jan Ramon (INRIA)

<b>Deliverable nature</b>	<b>Report</b>
<b>Dissemination level</b>	PUB
<b>Delivery date</b>	30-04-2025 [M24]
<b>Version</b>	1.0
<b>Total number of pages</b>	72
<b>Keywords</b>	GDPR, EHDS, DGA, cybersecurity, privacy, health data

## EXECUTIVE SUMMARY

The Legal and Ethics Compliance Report (D6.1) identifies the core legal, ethical and security requirements for the execution of the FLUTE project and paves the way to their operational implementation. It also summarizes the efforts made to ensure the project's compliance with these requirements.

The legal and ethical considerations are categorized into the following main sections:

1. Privacy and data protection (GDPR)
2. Appropriate safeguards under art. 89 GDPR
3. Synthetic data
4. Responsible and Trustworthy AI
5. Cybersecurity Laws
6. European Health Data Space (EHDS)
7. Data Governance Act (DGA)
8. Medical Device Regulation (MDR)

These requirements were first identified within the initial 12 months of the project and shared with FLUTE partners through the first version of this report. Each requirement was thoroughly explained, accompanied by guidance on its application within the project (referred to as the 'Resulting Requirements for FLUTE'). This guidance also outlined the measures taken to ensure compliance with relevant legal and ethical obligations.

Throughout the project, particular focus was placed on respecting legal and ethical principles, especially during the design of the FLUTE platform and the collection and use of medical data. Looking ahead, WP6 will continue working closely with partners to maintain alignment with evolving legal and ethical frameworks, including the European Health Data Space (EHDS). We will also focus on investigating legal requirements related to health data sharing in the for the purpose of secure federated learning, in view of leading and contributing to D6.2 GDPR guidelines and regulatory liaising. In particular, we will analyse the organizational measures that operators of federated learning platforms and data hubs should put in place to achieve and demonstrate regulatory compliance under GDPR, DGA, EHDS.

## DOCUMENT INFORMATION

Grant agreement No.	101095382	Acronym	FLUTE
Full title	<b>Federate Learning and mUlti-party computation Techniques for prostatE cancer</b>		
Call	HORIZON-HLTH-2022-IND-13-02		
Project URL	<a href="https://cordis.europa.eu/project/id/101095382">https://cordis.europa.eu/project/id/101095382</a> <a href="https://www.fluteproject.eu/">https://www.fluteproject.eu/</a>		
EU project officer	Nihal YILDIRIM		

Deliverable	Number	D6.1	Title	Legal and ethics compliance report
Work package	Number	WP6	Title	Ethics, regulatory acceptability, GDPR guidelines and standards
Task	Number	T6.1	Title	Legal and ethical compliance of the FLUTE platform

Date of delivery	Contractual	M24	Actual	M24
Status	version 1.0.	<input checked="" type="checkbox"/> Final version		
Nature	<input checked="" type="checkbox"/> R <input type="checkbox"/> DEM <input type="checkbox"/> DMP <input type="checkbox"/> DEC <input type="checkbox"/> ETHICS <input type="checkbox"/> OTHER			
Dissemination level	<input type="checkbox"/> Public <input checked="" type="checkbox"/> Sensitive			

Authors (partners)	TLX
Responsible author	Magdalena Kogut-Czarkowska
	magdalena.kogut@timelex.eu

Summary (for dissemination)	<p>The Legal and Ethics Compliance Report (D6.1) outlines the key legal, ethical, and security requirements guiding the FLUTE project, covering areas such as GDPR, synthetic data, trustworthy AI, cybersecurity, EHDS, DGA, and MDR. Identified in the project's first year, these requirements were shared with partners alongside practical guidance for implementation. The report highlights the project's ongoing commitment to legal and ethical integrity, particularly in platform design and medical data use.</p>
Keywords	GDPR, EHDS, DGA, GDPR, EHDS, DGA, MDR, cybersecurity, privacy, health data

<b>VERSION LONG</b>			
<b>Issue Date</b>	<b>Rev. No.</b>	<b>Author</b>	<b>Change</b>
<b>15-11-2023</b>	0.01	Magdalena Kogut-Czarkowska	Index
<b>17-01-2024</b>	0.02	Magdalena Kogut-Czarkowska	First internal draft
<b>29-04-2024</b>	0.1	Magdalena Kogut-Czarkowska	First draft including chapters on EHDS and DGA
<b>17-02-2025</b>	0.2	Magdalena Kogut-Czarkowska, Eleni Moraiti	First updated draft including chapters on AIA and cybersecurity certification
<b>26-03-2025</b>	0.3	Magdalena Kogut-Czarkowska, Eleni Moraiti, Marta Wilińska	Additional checks and format
<b>03-03-2025</b>	0.4	Magdalena Kogut-Czarkowska, Eleni Moraiti, Marta Wilińska, Jos Dumortier	Internal review
<b>04-03-2025</b>	0.5	Magdalena Kogut-Czarkowska, Marta Wilińska	Second updated draft for additional input from consortium partners
<b>14-04-2025</b>	0.6	Magdalena Kogut-Czarkowska	Document ready for peer-review
<b>28-04-2025</b>	1.0	Magdalena Kogut-Czarkowska, Marta Wilińska	Final version for submission

## Table of contents

<b>1. Introduction</b>	<b>9</b>
1.1. Purpose and scope	9
1.2. Overview of the FLUTE project	9
<b>2. Privacy and data protection (GDPR)</b>	<b>11</b>
2.1. Definition of personal data, anonymisation and pseudonymisation	11
2.2. Roles and responsibilities in relation to personal data	16
2.3. Compliance with data protection principles	17
2.4. Security of data and continuous risk assessment	20
2.5. DPIA requirements	21
2.6. Resulting requirements for FLUTE	22
2.7. Respect for data subjects' rights	22
2.8. Transfer of personal data	23
<b>3. Synthetic data</b>	<b>24</b>
3.1. Overview	24
3.2. Benefits of using synthetic data	24
3.3. Synthetic data in the guidance of Data Protection Authorities (DPAs)	25
3.4. Applicability of the GDPR to synthetic data	26
3.5. Resulting requirements for FLUTE	29
<b>4. Appropriate safeguards under art. 89 GDPR</b>	<b>30</b>
4.1. Overview	30
4.2. Resulting requirements for FLUTE	30
<b>5. Regulation on Artificial Intelligence (AI)</b>	<b>31</b>
5.1. Responsible and trustworthy AI	31
5.2. Overview of the EU Artificial Intelligence Act (AIA)	32
5.3. Limitations of AIA and risk-based approach	33
5.4. High-Risk AI Systems (HRAIS)	34
5.5. Requirements for HRAIS	35
5.6. HRAIS and harmonised standards	41
5.7. Application of the AIA	41
5.8. AIA and scientific research exceptions	41
5.9. Resulting requirements for FLUTE	41
<b>6. Cybersecurity laws</b>	<b>44</b>
6.1. NIS2	44
6.2. Cyber Resilience Act	45
6.3. Resulting requirements for FLUTE	47
6.4. Cybersecurity Certification	48
<b>7. European Health Data Space</b>	<b>52</b>
7.1. Overview	52
7.2. Key terms in EHDSR	52
7.3. Access to data for secondary use according to the EHDS	53
7.4. Data quality and utility labels	55
7.5. Secure processing environments	55
7.6. Resulting requirements for FLUTE	56

<b>8. Data Governance Act .....</b>	<b>57</b>
8.1. General information .....	57
8.2. Data Intermediation Services – definition and types.....	57
8.3. Requirements for providing DIS.....	59
8.4. Data altruism under the DGA.....	61
8.5. Resulting requirements for FLUTE.....	63
<b>9. Medical device regulation .....</b>	<b>64</b>
9.1. Software as a Medical Device.....	64
9.2. Resulting requirements for FLUTE.....	66
<b>10. Conclusions .....</b>	<b>68</b>
<b>11. References.....</b>	<b>69</b>

## Table of Figures

Figure 1 Summary of access to EHD for secondary purposes .....	54
Figure 2 Data intermediation services provider recognised in the Union label .....	61
Figure 3 Data altruism organisation recognised in the Union labels .....	62
Figure 4 Qualification of MDSW .....	65

## List of Tables

Table 1 Summary of guidelines issued by the EDPS, AEPD and ICO (UK DPA).....	25
Table 2 Risk Based approach in the AIA.....	33
Table 3 HRAIs.....	34
Table 4 Requirements for HRAIs.....	36
Table 5 Categories of DIS under DGA .....	58
Table 6 The main differences between DISP and RDAO .....	63

## ABBREVIATIONS AND ACRONYMS

**AEPD:** Spanish Data Protection Authority

**AI:** Artificial Intelligence

**AIA:** EU Artificial Intelligence Act

**CB:** Accredited Conformity Assessment Body

**CJEU:** Court of Justice of the European Union

**CNIL:** French Data Protection Authority

**CRA:** Cyber Resilience Act

**CsPCa:** clinically significant Prostate Cancer

**CSIRT:** Computer security incident response team

**DGA:** Data Governance Act

**DIS:** Data Intermediation Service

**DISP:** Data Intermediation Service Provider

**DoA:** Description of Action

**DPA:** Data Protection Authority

**DSR:** Data subject rights

**EC:** European Commission

**EDPB:** European Data Protection Board

**EDPS:** European Data Protection Supervisor

**EHDS:** European Health Data Space

**ENISA:** European Union Agency for Cybersecurity

**EU:** European Union

**FL:** Federated Learning

**FMM:** Federated Model Management

**GDPR:** General Data Protection Regulation

**HRAI:** High-risk AI system

**ICT:** Information Communication Technology

**IP:** Intellectual Property

**IVDR:** In Vitro Medical Device Regulation

**JCA:** Joint Controllorship Arrangement

**MDCG:** Medical Device Coordination Group

**MDR:** Medical Device Regulation

**MDSW:** Medical Device Software

**OJEU:** Official Journal of the European Union

**PCa:** Prostate Cancer

**PET:** Privacy Enhancing Technology

**PIMS:** Personal Information Management Systems

**RDAO:** Recognized Data Altruism Organization

**SME:** Small and Medium Enterprise

## 1. Introduction

### 1.1. Purpose and scope

The purpose of this deliverable D6.1 is to outline the legal and ethical requirements for the project and to provide an overview of the legal and ethical support provided to the consortium throughout the implementation of the project.

The document builds on the D5.1 of the TRUMPET project and the requirements which were formulated there, expanding onto new, additional points, which are outlined below.

The initial version of the deliverable was first presented to FLUTE partners in M12 of the project, identifying the core legal, ethical and security requirements for the execution of the FLUTE. The document was further developed and detailed, including by operationalising the objectives in cooperation with other tasks and work packages in FLUTE.

As a final version, the document presents the key legal and ethical issues across the following domains:

1. Privacy and data protection (GDPR)
2. Appropriate safeguards under art. 89 GDPR
3. Synthetic data
4. Responsible and Trustworthy AI
5. Cybersecurity Laws
6. European Health Data Space (EHDS)
7. Data Governance Act (DGA)
8. Medical Device Regulation (MDR)

For each of these topics, this deliverable provides an overview and explanation, which will be linked to specific requirements for the FLUTE project, and information about their implementation in the current state of the project.

### 1.2. Overview of the FLUTE project

The Project has the following objectives:

- advance and scale up data-driven healthcare by developing novel methods for privacy-preserving cross-border utilization of data hubs,
- develop a novel federated AI toolset for diagnosis of clinically significant prostate cancer and perform a multi-national clinical validation study of its efficacy, which will help to improve predictions of aggressive prostate cancer while avoiding unnecessary biopsies, thus improving the welfare of patients and significantly reducing the associated costs,
- create a privacy-enforcing platform that will provide innovators with a provenly secure environment for federated healthcare artificial intelligence (AI) solution development, testing and deployment, including the integration of real-world health data from the data hubs and the generation and utilization of synthetic data,

- contribute to the global HL7 FHIR standard development and create novel guidelines for GDPR-compliant cross-border Federated Learning in healthcare.

In this regard, FLUTE will develop and validate a flexible and highly scalable open-source platform that will intermediate between the privacy-sensitive data hubs and cross-border users/innovators, with the data hubs hosting local learning nodes connected to a central federated model aggregator in a FL setting.

The data used in the context of the use case stem from medical centres (Data Providers) located in 3 European countries (IRST, CHU, VHIR) while SR enlarges and geographically diversifies FLUTE's data sets by providing retrospective data stemming from proprietary data collections collected from collaborating clinics.

The FLUTE platform will provide the privacy guarantees and scalability the lack of which prevented both the development and the practical deployment of such multi-centres, cross-border healthcare AI models so far. The FLUTE platform will be piloted and validated in the a use case - Robust prostate cancer prediction using FL across borders of 4 European countries.

The main results expected from FLUTE's research work are:

- achieve robust performance in clinically significant Prostate Cancer (csPCa) cancer detection, considering variations in the prevalence of the disease in different regions of Europe,
- develop and validate a cross-border federated AI solution on the FLUTE platform for the diagnosis of csPCa in different regions of Europe,
- provide novel intermediating FLUTE platform that will make the siloed datasets available to researchers under privacy protection that complies with and exceeds GDPR requirements, in standard FL settings and AI model development workflows,
- ensure thorough validation in one research use case,
- validate the Barcelona Risk Calculator models (BCN-RC 1 and 2) with the inclusion of imaging biomarkers extracted from mpMRI/bpMRI for the detection of csPCa with data from multiple regions,
- generate synthetic clinical data and images to train convolutional neural network models without the risk of re-identifying sensitive data, ensuring privacy preservation.

## 2. Privacy and data protection (GDPR)

### 2.1. Definition of personal data, anonymisation and pseudonymisation

#### 2.1.1. Overview

The provisions of the General Data Protection Regulation<sup>1</sup> (GDPR) outline the requirements for fair and lawful processing of personal data. To determine whether these requirements apply, the starting point must always be the definition of ‘processing of personal data’. The term processing encompasses any operation on the data, even storing data or consulting it. This very broad definition covers any use of personal data. In particular, it covers cases where AI is trained on personal data, and when AI is used to analyse or reach decisions about individuals.

Personal data, on the other hand, is defined as any information relating to an identified or identifiable natural person (a ‘data subject’). The three main elements of the definition are: (i) any information; (ii) relating to a natural person; and (iii) with that person being identified or identifiable.

This means that any information linked to a person in any way is personal data as long as the person can either be directly identified or the information makes the person identifiable i.e., if you gather additional information, you are able to identify the person.

Information that directly identifies a person (e.g., the person’s name, ID number, etc.) is the most known and intuitively understood form of personal data. It is a common mistake to only consider these types of straightforward personal data as relevant when applying the GDPR.

However, information that does not directly identify a person, but requires additional information to be acquired, held or used to identify the person also constitutes personal data. An example is a phone number found on the Internet. The number does not directly identify a person, but it is usually possible to find other information on the Internet, call the number, or contact the mobile service provider to get additional information that allows you to identify its owner. That possibility makes the person identifiable (i.e., possible to be identified) and thus makes the initial piece of information personal data. This expansive scope of the meaning of ‘identifiable’ implies that the GDPR applies to all kinds of information that is linked or relates to a person, irrespective of whether that individual is currently identified, as long as it is theoretically possible to identify them.

The definition of personal data is connected with the concepts of anonymization and pseudonymization of personal data.

- **Anonymous information** is one which does not relate to an identified or identifiable natural person or relates to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly (recital 26 GDPR).
- **Pseudonymization** is the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational

---

<sup>1</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC

measures to ensure that the personal data is not attributed to an identified or identifiable natural person (Article 4(5) GDPR).

Data can only be considered anonymous when it no longer relates to an identified or identifiable natural person. This is different from pseudonymization, which aims to make it much *more difficult* for a person to be identified but does not altogether remove the possibility of identification. In other words, anonymization must make it impossible under ‘the means reasonably likely to be used’ for the controller or another third party to use the data to identify a natural person and this process must be irreversible. Therefore, if personal data is properly anonymised, the GDPR does not apply. In contrast, the GDPR always remains applicable whenever data is only pseudonymised. This is the main difference from a legal point of view between anonymisation and pseudonymisation.

In practice, the de-alienation of pseudonymized data from anonymous data is difficult. There is lack of clear standards on how personal data, including medical data, should be anonymized. The European Data Protection Board (EDPB, an administrative body that under the GDPR provides guidance on the interpretation of the GDPR) did not issue guidelines on anonymization under the GDPR. Even more importantly, there are conflicting views of EU Data Protection Authorities (DPAs) and the courts on the matter. In essence, the approach of the DPAs is stricter than the ‘risk based’ or contextual approach followed by the EU courts<sup>2</sup>.

The DPAs refer mostly to Opinion 05/2014 on Anonymisation Techniques<sup>3</sup>, which was issued in 2014 by Working Party Article 29 (an advisory body established on the previous data protection directive). Opinion 05/2014 states that ‘identification’ means not only the possibility of retrieving a person’s name and/or address but also includes potential identifiability by singling out (singularisation), linkability and inference. Those are understood as follows:

- **Singularisation:** the possibility of extracting from a data set some records (or all records) that identify a person.
- **Linkability:** the ability to link at least two records of a single data subject or a group of data subjects, either in the same database or in two different databases. If the attacker can determine (e.g. by correlation analysis) that two records are assigned to the same group of persons but cannot single out the persons in this group, then the technique is resistant to singling out, but not to linkability.
- **Inference:** the possibility of deducing with significant probability the value of an attribute from the values of a set of other attributes. The second step would be to apply the set of anonymisation techniques that eliminate this risk or reduce it as much as possible (more details are in Annex to the Opinion).

Moreover, Opinion 05/2014 states that ‘*means likely reasonably to be used to determine whether a person is identifiable*’ are those to be used ‘*by the controller or by any other person*’. This led to the conclusion of the Opinion 05/2014 that ‘*when a data controller does not delete the original (identifiable) data at event-level, and the data controller hands over part of this dataset (for example after removal or masking of identifiable data), the resulting dataset is still personal data.*’ Critiques argue that Opinion 05/2014 is not sufficiently nuanced, as under the GDPR account should be paid to broader circumstances in which the

<sup>2</sup> Weitzenboeck, Emily M., et al. "The GDPR and unstructured data: is anonymization possible?" *International Data Privacy Law* 12.3 (2022): 184-206.

<sup>3</sup> Article 29 Working Party, Opinion 5/2014 on Anonymization Techniques. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

data set is shared, which are case specific and need to take into account including cost of and time required for identification, and the available technology.

Recently, the EDPB adopted new guidelines on pseudonymisation<sup>4</sup> (public consultation ongoing). These guidelines do not explicitly clarify whether pseudonymized data should automatically be classified as personal data for all parties involved, simply because the data subjects remain identifiable due to the continued existence of information that enables their identification [as reflected in the position of the European Data Protection Supervisor (EDPS) and the EDPB in the SRB v. EDPB case]. However, certain sections of the guidelines (highlighted below) may imply that this is the general stance of the EDPB regarding pseudonymized data. In particular:

*22. Pseudonymised data, which could be attributed to a natural person by the use of additional information, is to be considered information on an identifiable natural person, and is therefore personal. This statement also holds true if pseudonymised data and additional information are not in the hands of the same person. If pseudonymised data and additional information could be combined having regard to the means reasonably likely to be used by the controller or by another person, then the pseudonymised data is personal. Even if all additional information retained by the pseudonymising controller has been erased, the pseudonymised data becomes anonymous only if the conditions for anonymity are met.*

*49. Note that pseudonymisation by the original controller also aids controllers who are recipients of pseudonymised data in fulfilling their data protection obligations, in particular with regard to the data minimisation principle, data protection by default and the maintenance of an appropriate level of security.'*

In turn, the European Courts did not refer to the criteria of Opinion 05/2014 in their case law. The most well-known verdict on the topic is the *Breyer* decision<sup>5</sup> from the Court of Justice of the European Union (CJEU). In that case, the CJEU explicitly noted the importance of context when assessing the 'means likely reasonably to be used' to identify the data subject. The Court said specifically that it must be determined '*whether the possibility to combine a dynamic IP address with the additional data held by the internet service provider constitutes a means likely reasonably to be used to identify the data subject. Thus, as the Advocate General stated essentially in point 68 of his Opinion, that would not be the case if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant*'. In simpler terms: in this particular context, the CJEU decided that a dynamic IP address should be considered personal data in the processing operations of the website operators because they could link the IP addresses to a natural person with the cooperation of internet service providers; such cooperation would be legally permissible, under certain circumstances.

In a more recent case (T-557/20<sup>6</sup>), the General Court (GC) was considering how to qualify information shared between a controller (SRB), who was sharing pseudonymized data with a recipient (Deloitte), who did not have access to codes. Data protection authority relevant in this case (EDPS) was arguing that:

<sup>4</sup> EDPB Guidelines 01/2025 on Pseudonymisation, Available at: [https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation\\_en](https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation_en)

<sup>5</sup> CJEU Judgement of 19/10/2016, Case C 582/14, Patrick Breyer v Bundesrepublik Deutschland, ECLI:EU:C:2016:779

<sup>6</sup> General Court (Eighth Chamber, Extended Composition), Judgment of 26/04/2023, Case T 557/20, Single Resolution Board (SRB) v European Data Protection Supervisor (EDPS), ECLI:EU:T:2023:219

*‘the data the SRB shared with Deloitte were pseudonymous data, both because the comments in [the consultation phase] were personal data and because the SRB shared the alphanumeric code that allows linking the replies given in [the registration phase] with the ones given in [the consultation phase] – notwithstanding the fact that the data provided by the participants to identify themselves in [the registration phase] were not disclosed to Deloitte. [...] The EDPS finds that Deloitte was a recipient of the complainants’ personal data’.*

However, the GC disagreed with this position. The Court noted that:

*‘in order to determine whether the information transmitted to Deloitte constituted personal data, it is necessary to put oneself in Deloitte’s position in order to determine whether the information transmitted to it relates to ‘identifiable persons’. [...] Deloitte’s situation can be compared to that of the online media services provider referred to in [Breyer] [...]. The SRB’s situation can also be compared to that of the internet service provider in that case [...]. Therefore, pursuant to paragraph 44 of the judgment of [Breyer], it was for the EDPS to examine whether the comments transmitted to Deloitte constituted personal data for Deloitte.’*

Thus, similarly as in Breyer, the Court puts emphasis on the means of identification available from the position of the recipient of the information and did not refer to Opinion 05/2014. While the SRB case is subject to appeal and the GC did not take a position on the status of the data at hand (it merely observed that the situation of Deloitte was not properly assessed by the EDPS), it clearly shows that the courts present a different view on the assessment of anonymity of data than the DPAs. Still, as the DPAs are the authorities which are empowered to enforce the GDPR, their position cannot be disregarded. Intervention of the DPA may lead to administrative liability (including fines) and reputational damage to the controllers; civil liability may also follow.

Following the EDPS’ appeal on this judgement the Advocate General has issued his opinion<sup>7</sup> on the subject. The Advocate General stated that obligations arising from the processing of personal data should always be assessed from the perspective of the entity concerned. Pseudonymized data should not automatically be deemed personal data for all parties involved merely because the data subjects remain identifiable due to the continued existence of information enabling their identification. Instead, pseudonymized data should be considered personal data only for the parties that can reasonably identify the data subject. Notwithstanding the above, the Advocate General emphasizes that, in this case, the obligation to inform data subject of the recipients of their data, as required under Article 15(1)(d) of Regulation 2018/1725, is part of the relationship between the data subjects and the controller. Since this obligation arises at the time of collection, before data is pseudonymized, the controller should disclose the identity of the data recipients regardless of whether the data will constitute personal data from the recipient’s perspective. According to the Advocate General the purpose of the obligation to provide information concerns the relationship between the controller and the data subjects and is intended to enable the latter to give their informed consent before the transfer.

### **2.1.2. Implementation of the requirements in FLUTE**

Under GDPR, data protection requirements are applicable to all processing operations in the project whenever not fully anonymised data is being processed. In particular:

---

<sup>7</sup>Opinion of Advocate General Spielmann delivered on 6/02/2025, Case C 413/23 P European Data Protection Supervisor v Single Resolution Board, Available at <https://curia.europa.eu/juris/document/document.jsf?text=&docid=295078&pageIndex=0&doclang=EN>

- Project partners should provide and agree upon the anonymisation and pseudonymisation measures to be followed, acknowledging the difficulties in ‘fully anonymizing’ personal data.
- Ultimately, out of caution, FLUTE may need to treat all datasets as personal data, as the Clinical partners (also referred to in other project documents as ‘Data Owners’ or ‘Data Providers’) have the original data of the patients.
- Clinical partners should obtain ethics approvals from their ethical committees and are responsible for the personal data de-identification.
- Clinical partners should keep the training datasets in a dedicated, safe infrastructure (data marts), separate from the clinical records. The data should be pseudonymized and de-identified under the agreed upon measures. Pseudonyms should be securely stored and kept separately.
- The teams which pseudonymize the data should not be the same which then use the data for research purposes (factual separation) or an automatized pseudonymization process should be implemented.
- Controllers should determine whether anyone would have the motivation to carry out re-identification and, whether the re-identification may/is likely to be successful, making use of the ‘motivated intruder’ test<sup>8</sup>. The attacker is perceived as a motivated and competent third party with access to resources commensurate with the motivation it may have for the re-identification. This has relevance to the threat models defined in the FLUTE project that rely on two types of attackers, namely ‘curious but honest participants’ and ‘dishonest and colluding participants’ (T2.4). In these scenarios, the attacker successfully assumes a man-in-the-middle position in the communication channel between the data hubs and the FLUTE platform or if the attacker compromises the Federated Model Management (FMM), turning it into a ‘curious but honest aggregator’ that honestly performs model aggregation but inspects the model updates and uses them to re-identify subjects. In assessing what level of anonymisation is necessary, all methods reasonably likely to be used by someone (either an ‘intruder’ or an ‘insider’) should be considered to identify an individual data subject given the current state of technology and the information available to such a person at present.

To assist FLUTE partners in complying with the legal requirements in the data protection field, dedicated WP6 actions were taken to explain and operationalize the requirements in collaboration with the clinical partners and technical partners involved in the development of the platform. In particular:

- GDPR workshop, organised by TLX, saw the participation of all FLUTE partners, focusing on key areas such as the principles of data protection and GDPR, defining roles and key responsibilities under the GDPR, differentiating between anonymisation and pseudonymisation, the data protection impact assessment (DPIA), and outlining subsequent steps necessary for achieving GDPR compliance, such as how the GDPR requirements translate into stakeholder and user requirements.
- Workshop ‘Understanding the notion of personal data’ organised by TLX, focused on explaining the scope of personal data as interpreted by the case law of the CJEU. The workshop followed an evolutionary approach to case law and emphasized on the different nuances that stem from the contextual assessment that the CJEU regularly applies. The main contribution of the workshop is that it helped partners understand that there are no one-size-fits-all answers but a careful examination of

---

<sup>8</sup> WP29, Opinion 6/2013 on open data and public sector information (‘PSI’) reuse. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf)

the circumstances at hand is always required to assess whether a processing operation involves personal data.

- TLX has compiled various topics of interest to the partners and has provided comments and feedback to several technical deliverables, and promoted dialogue among legal and technical stakeholders to clarify the technical partners' inquiries and shed light on various pertinent subjects from a legal standpoint. Also, TLX actively participated in several WP calls by providing legal and/or ethical assistance in specific questions arising during the development of the FLUTE platform.
- TLX is committed to contributing to publications and further investigation of the notion of personal data during the remainder of the project.

## 2.2. Roles and responsibilities in relation to personal data

### 2.2.1. Overview

The GDPR defines two main concepts for entities dealing with personal data:

- The controller, which is the entity that determines the purposes and means of the processing (i.e., why the data is being processed, what data, about which type of data subjects, in what way, etc.).
- The processor, which processes data on behalf of the controller and receives instructions as to what data to process, for which purposes and how (even though technical decisions may be made by the processor as well).

The **controller** is the responsible entity under the GDPR which must ensure that all processing of personal data is done in compliance with the law. This includes amongst others:

- applying the GDPR principles relating to the processing of personal data;
- carrying out a data protection impact assessment prior to processing (where needed);
- providing information to the data subjects and enabling the exercise of data subject rights;
- respecting all other GDPR obligations (technical and organizational measures).

If the controller uses any **processors** to implement its processing, it must have a contract with that processor (data processing agreement) which complies with Article 28 GDPR.

In addition, whenever two or more controllers together decide on the 'why' and 'how' of the processing, a situation of **joint controllership** may exist. Joint controllership in a research project implies setting up common structures, through which processing of personal data will be done for specific common goals, and common or converging decisions on these elements by the involved parties. It typically implies an element of operational co-decision and cooperation. The EDPB Guidelines 07/2020<sup>9</sup> include an example:

*'Several research institutes decide to participate in a specific joint research project and to use to that end the existing platform of one of the institutes involved in the project. Each institute feeds personal data it already holds into the platform for the purpose of the joint research and uses the data provided by others through the platform for carrying out the research. In this case, all institutes qualify as joint controllers for the personal data processing that is done by storing and disclosing information from this platform since they have decided together the purpose of the processing and the means to be used (the existing platform). Each of the institutes however is a separate controller*

<sup>9</sup> EDPB Guidelines 07/2020, pg. 24.

*for any other processing that may be carried out outside the platform for their respective purposes.’*

As noted by the guidelines, joint participation can take the form of a *common decision* taken by two or more entities or result from *converging decisions* by two or more entities, where the decisions complement each other and are necessary for the processing to take place in such a manner that they have a tangible impact on the determination of the purposes and means of the processing. An important criterion is that the processing would not be possible without both parties’ participation in the sense that the processing by each party is inseparable, i.e. inextricably linked.<sup>10</sup> Joint controllers should determine and agree on their respective responsibilities for compliance with the obligations under the GDPR in a binding document.<sup>11</sup> The existence of joint responsibility does not necessarily imply equal responsibility of the various operators involved in the processing of personal data. The CJEU in its rulings<sup>12</sup> clarified that those operators may be involved at different stages of that processing and to different degrees so that the level of responsibility of each of them must be assessed with regard to all the relevant circumstances of the particular case.

Determining who are the responsible entities and whether there is a situation of joint controllership is especially relevant for the FLUTE platform. While the FLUTE platform may be designed with input from all partners, whether this will imply a situation of joint control depends on how the platform will process personal data operationally and who will be in charge of its operation. This does not however imply that all the partners of the FLUTE consortium must assume such a role. Certain partners may not make any decisions regarding personal data nor process such data and thus would not have a role defined under the GDPR.

## 2.2.2. Resulting requirements for FLUTE

Given the above, in order to impose the proper structure of data protection framework in the FLUTE project it was crucial to carry out an assessment of roles of all consortium partners.

- This assessment was made on a factual, rather than a formal, analysis of their actual influence with respect to determining the purposes and means of the collecting the data, their storage and use for the purposes of the project. In FLUTE this work started with a data mapping document, which outlined the following roles of the partners in the data processing. Following detailed discussions, it was concluded that FLUTE partners act as joint controllers, however they do not have the same role and obligations.
- Accordingly, the legal partner led the development of a joint controller agreement (JCA), which was prepared as an initial draft for signature in April 2024 and then was appropriately filled in and subjected to a thorough review by all partners. This important document outlines the roles and responsibilities of all partners regarding clinical data management. The document has been signed by consortium partners. Moreover, following the inclusion of Siemens Romania in the consortium, TLX prepared a separate amendment for their inclusion in the JCA.

## 2.3. Compliance with data protection principles

### 2.3.1. Overview

The GDPR implies compliance with seven key principles:

<sup>10</sup> EDPB Guidelines 07/2020, pg. 3.

<sup>11</sup> Article 26 GDPR.

<sup>12</sup> Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie, (C- 210/16), Tietosuojavaltutettu v Jehovan todistajat — uskonnollinen yhdyskunta (C-25/17), Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW eV (C-40/17).

1. **lawfulness, fairness and transparency** – meaning that a legal basis for any data processing (including but not limited to consent) must be available, and that the persons concerned must be appropriately informed of how their data will be used;
2. **purpose limitation** – meaning that data must be collected for specific purposes, and may thereafter only be used for compatible purposes;
3. **data minimisation** – meaning that data collected and used in the project must be as minimal as possible, taking into account the intended purposes;
4. **accuracy** – meaning that measures must be taken to ensure the quality and accuracy of the data, and that measures must be available to detect and remedy problems;
5. **storage limitation** – meaning that data may only be retained for as long as necessary given the intended purposes, and that it must thereafter be deleted or anonymised;
6. **integrity and confidentiality** – meaning that data must be protected by appropriate technical and organisational measures to ensure its confidentiality, integrity and availability;
7. **accountability** – meaning that responsible entities must be identified, and that appropriate controls (such as logs) are available to ensure that any problems can be attributed to the correct entity.

A detailed discussion of all the principles is beyond the scope of this report, as this was explained in D5.1 in TRUMPET and FLUTE builds on the work done under that project. Below we discuss the key considerations for FLUTE Project set up.

### 2.3.2. Resulting requirements for FLUTE

#### 2.3.2.1. Lawfulness

FLUTE will re-use retrospective data for the case studies. Personal data processing within FLUTE may take place only if there is an applicable legal basis for the processing:

- Data Providers are responsible for meeting this requirement. Each Data Provider must check whether they can assure that they have the appropriate legal basis to contribute Data to FLUTE. In particular, it should be assessed by the Data Providers whether the research activities under the project may be covered under the consent already given by the patients upon their admittance to the hospitals and/or whether national exceptions for consent for secondary use of data for scientific research purposes apply (see section Purpose Limitation). If this legal basis is not available, only anonymized data should be submitted to the project.
- If any patients take part in prospective data collection, they must provide consent for the participation and the processing of data.
- Although FLUTE will require the Data Providers to confirm that they have a legal basis for the processing of personal data, the actual responsibility of meeting ethics and legal requirements remains with the Data Providers.

#### 2.3.2.2. Fairness and transparency

The rule of transparency states that data must be processed in a transparent manner in relation to the data subject. This rule is closely linked to information rights (Art. 13-14 GDPR) and data subject rights ('DSR', Art. 15-21 GDPR). Under those provisions, in general, data subjects must be properly informed about the details of processing of their data for a specific purpose (also in case the data is not collected

directly from the data subject). Detailed list of information to be provided is included in Articles 13 and 14 GDPR. There are very limited exceptions from this requirement.

- In the context of FLUTE, unless an exception applies, the transparency of the processing vis-à-vis the data subjects should be ensured through appropriate notices and policies, such as properly drafted information made available to patients in their respective languages (unless exception applies), privacy notice on the repository website and clear description in any agreements to be entered into by the Data Providers, such as the FLUTE JCA.

### **2.3.2.3. Purpose limitation**

Under this rule, the Project partners may only collect and use data for a ‘specified, explicit, and legitimate’ purpose. The use of the data for another, ‘incompatible’ or unrelated purpose is not allowed. However, there are exceptions allowing ‘further processing’, e.g., for research purposes and statistical purposes. If any data is planned to be re used in the context of FLUTE, but was collected for a different purpose, it needs to be assessed whether such re-use will be permitted. This assessment should take into account, in particular:

- any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
- the context in which the personal data have been collected, such as the information provided to the data subjects;
- the nature of the personal data, in particular if sensitive data (such as health data) is involved;
- the possible consequences of the intended further processing for data subjects;
- the existence of appropriate safeguards, which may include encryption or pseudonymisation.

### **2.3.2.4. Data minimisation**

Under the rule of data minimalization data collected and used in the FLUTE project must be as minimal as possible, taking into account the intended purposes. This implies that data should not be held for further use, unless this is essential for reasons that were stated in advance. Furthermore, data should be de-identified (anonymized or pseudonymized) and should only include as much information as required to successfully answer the research question(s). Data minimization should also be observed by restricting the access to data on a ‘need to know’ basis and limiting the copying of data, whenever possible.

In FLUTE, the discussions about data minimization were held in the context of both:

- scope of data processed: during the definition of the study protocol and determination of the data fields which will be used in the project.
- manner of access: As a general rule, FLUTE project starts from the idea that sensitive data must not leave the premises of the Data Providers. Thus, once the data is prepared by the hospital, it resides on a server (called ‘Data Provider Node’) protected by the hospital’s own infrastructure. This server then exchanges encrypted messages with other Data Providers via FLUTE platform to collaboratively compute aggregates (e.g., averages of attributes or gradients, or other statistics). Limited exceptions to this rule will be defined in the JCA .

Data minimisation is one of the guiding principles in the design of FLUTE platform.

### **2.3.2.5. Accuracy**

According to the accuracy principle, personal data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay. The principle of accuracy is closely linked to the quality of data.

- FLUTE must take and document steps to ensure that medical data is accurate, of appropriate quality and contain accurate metadata. For data, which is pseudonymized, there should also be a process of removing and correcting incorrect or misleading data, in particular at the legitimate request of a data subject (unless an exception applies). Hence, if the data is pseudonymized (not anonymized) there should a possibility to reverse that process and correct the data. However the possibility and necessity of such an action should be determined on case by case bases.

### **2.3.2.6. Storage limitation**

Under the storage limitation principle, personal data must be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed. However, personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) GDPR subject to implementation of the appropriate technical and organisational measures required by the GDPR in order to safeguard the rights and freedoms of the data subject. This does not, however, imply that data stored for scientific research may be kept forever. Data subjects must be informed about the retention period, or at least its basis and rationale. After this period lapses, the data must be removed or anonymized.

Data Providers and other controllers within the FLUTE project must define a retention period for the pseudonymized data that is collected for the participation in FLUTE and/or available in the repository. The specific timeframes should be determined in accordance with national laws.

### **2.3.2.7. Integrity and confidentiality**

The obligation described in this point means that personal data must be protected by appropriate technical and organisational measures to ensure its confidentiality, integrity and availability, as provided in Article 32 GDPR. This is closely related to requirements to provide appropriate safeguards under Article 89(2) GDPR. As this is a crucial part of the project, this will be described in more detail in point 2.4 below.

### **2.3.2.8. Accountability**

Data controllers within FLUTE must be able to demonstrate compliance with the GDPR principles. One of the documents that serves to fulfil this requirement is this deliverable.

## **2.4. Security of data and continuous risk assessment**

### **2.4.1. Overview**

Article 32 GDPR concerns the security of the processing. The controller must implement appropriate technical and organizational measures to ensure an appropriate level of security according to the risks posed by the processing. For this, the state of the art must be considered, as well as the costs of implementation and the nature, scope, context and purposes of processing, including the risk of varying likelihood and severity for the rights and freedoms of natural persons. Measures that can be taken include:

- the pseudonymization and encryption of personal data;

- the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
- the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
- a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.

If the controller hires a processor, such processor must provide guarantees to implement security measures (Article 28(3) GDPR). The processor is responsible to take all measures pursuant to Article 32 and may be held liable for infringement of this obligation. Whether a security level is appropriate depends on the precise risks presented by a given processing operation. These can range from accidental or unlawful destruction, loss and alteration to unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed.

Controllers and processors must ensure their employees handling personal data do not process them except as instructed.

#### **2.4.2. Resulting requirements for FLUTE**

FLUTE should document the security choices based on the risks to personal data and implement appropriate safeguards.

- The description of those safeguards should be made available to Data Providers. In turn, each Data Provider must check whether the safeguards offered by FLUTE fulfil their requirements. This description was included in FLUTE JCA.
- FLUTE medical and technical partners are responsible for inspecting whether the FLUTE Platform is appropriate for protecting the personal data stored in the local nodes at a level that is appropriate for the ethics and privacy constraints applicable, and for ensuring that personal data can be shared through the FLUTE platform in a way that does not run counter to the rights and interests of research participants. Data processing in FLUTE platform is, on the legal side, subject to data protection impact assessment (Section 2.5 below). From the technical side, apart from internal reviews, the privacy and security of the platform will be verified by penetration testing conducted by a third party.

### **2.5. DPIA requirements**

#### **2.5.1. Overview**

Article 35 GDPR requires a prior assessment of the impact of the envisaged processing operations on the protection of personal data, particularly when using new technologies and when the processing – taking into account its nature, scope, context and purposes – is likely to result in a high risk to the rights and freedoms of natural persons. The data protection officer may be requested advice.

According to the GDPR, a DPIA is required when the processing concerns an automated processing providing systematic and extensive evaluation of personal aspects relating to natural persons and on which decisions are based that produce legal effects concerning the natural person; when processing on a large-scale special categories of data; or when systematically monitoring a publicly accessible area on a large scale. A public list of these operations is maintained by the supervisory authority.

A DPIA must contain:

- a systematic description of the envisaged processing operations and the purposes of the processing;
- an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- an assessment of the risks to the rights and freedoms of data subjects; and
- the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance, taking into account the rights and legitimate interests of data subjects and other persons concerned.

A DPIA is not required if the processing is prescribed by law and if a general impact assessment was already conducted in adopting that legal basis.

If the processing can result in a high risk, the controller should contact the supervisory authority. If the authority finds that the processing would infringe upon the GDPR, or that the risks are insufficiently mitigated, it may provide a negative advice to the processing. Member States may allow authorities to require controllers to obtain prior authorization.

## 2.6. Resulting requirements for FLUTE

In the context of FLUTE, the project combines health data from many sources and thus processing on a large scale of special categories of data is likely to take place. However, FLUTE is not a single entity (controller). For this reason, a DPIA was performed with the participation of Data Providers (acting as controllers). They were asked to describe in detail the data processing operations of importance for the development and use of the FLUTE platform, provide information on the applicable technical and organisational measures to ensure the integrity of personal data and present the mitigating measures to address potential risks. The DPIA serves as an accountability tool and supports the compliance of the project with the GDPR requirements.

## 2.7. Respect for data subjects' rights

### 2.7.1. Overview

Where the project will use personal data, including pseudonymized data, data subject rights give the patient possibilities to obtain information about processing and - in cases specified by the law - influence such processing. These rights are set out in Chapter III of the GDPR, and include the rights of the data subject to:

- obtain information about the processing of their personal data (Article 15(1) GDPR)
- obtain access to the personal data held about them and copy of the data (Article 15(1) and 15(3) GDPR)
- ask for incorrect, inaccurate or incomplete personal data to be corrected (Article 16 GDPR);
- request that personal data be erased when it's no longer needed or if processing it is unlawful (right to be forgotten, Article 17 GDPR);
- request the restriction of the processing of their personal data in specific cases (Article 18 GDPR);
- receive personal data in a machine-readable format and send it to another controller ('data portability', Article 20 GDPR);

- object to the processing of their personal data for marketing purposes or on grounds relating to their particular situation (Article 21 GDPR);
- request that decisions based on automated processing concerning them or significantly affecting them and based on their personal data are made by natural persons, not only by computers. They also have the right in this case to express their point of view and to contest the decision (Article 22 GDPR).

However, according to Article 89(2) GDPR there are possible exceptions to the data subjects' rights. In particular, where personal data is processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from certain rights referred to above in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

### **2.7.2. Resulting requirements for FLUTE**

Within FLUTE pilots the health care professional must ensure that patient rights as defined under national law – notably access to the patient record, the right to a second opinion, the right for amendment or correction of a record, as applicable under national law – are respected at all times.

With respect to data made available on the platform (stored in the federated nodes), the respective controllers of this data must implement policies to ensure intake and execution of the patient rights with respect to their non-anonymized data.

## **2.8. Transfer of personal data**

### **2.8.1. Overview**

The GDPR includes an entire chapter on the transfer of personal data to a 'third country' (non-EEA country) or international organisation. Only when one of the 'transfer mechanisms' listed in articles 45 to 47 of the GDPR is complied with, the transfer of personal data to a third country or international organisation may take place.

Furthermore, the transfer of personal data to a third country or international organisation in itself constitutes a processing activity, which requires a legal ground as well as fulfilling the conditions related to processing special categories of data, namely medical image and health data.

In other countries (not part of the EU), the requirements for the transfer of personal data to third countries are often either inspired by the GDPR, or mainly put emphasis on the principles of transparency, purpose limitation and consent. Thus, any transfers to non-EEA partners need to be considered under the applicable rules.

### **2.8.2. Resulting requirements for FLUTE**

To the extent that personal data are processed in FLUTE, it should be ensured that all national requirements relating to a transfer of personal data to third countries or international organisations are observed. The FLUTE project includes partners from the EU, UK and Israel. However, during the JCA drafting, the partners from UK and Israel have indicated that they will not need to transfer or access the data. Should this change in the future, appropriate amendments should be concluded.

## 3. Synthetic data

### 3.1. Overview

Synthetic data does not have a legal definition. There are two understandings of this term in the legal literature and guidance:

- **Broader definition:** data artificially generated to resemble the characteristics of real data, including their structure and statistical distribution - see for example AEPD (Spanish Data Protection Authority).<sup>13</sup>
- **Narrow definition:** synthetic data is one generated by a mathematical model or algorithm (synthetic data generator). For example, EDPS notes that synthetic data is artificially generated from original data and a model trained to reproduce characteristics and structure of that data.<sup>14</sup>

Furthermore, the generation of synthetic data can take various forms, including its production from real datasets or its creation 'from scratch' by leveraging knowledge and expertise gathered by data analysts on specific dependencies. It can also result from a combination of these approaches, incorporating both real data and expert knowledge to create synthetic datasets.<sup>15</sup>

The primary objective of synthetic data is to preserve the characteristics and properties of real data tailored to a specific use case<sup>16</sup>. Notably, the determination of which properties of the real data should be preserved hinges on the intended purpose of the data usage. For instance, distinct data qualities are required when assessing the storage capacities of an IT system compared to using the data for training an AI model in cancer detection.

In certain applications, the relevance of data quality, in the sense of the close resemblance between synthetic data and real data, may be nonessential. For example, when synthetic data is used to train self-driving vehicles, the occurrence of risky situations in this dataset may need to be more frequent than in real life driving conditions.<sup>17</sup> Hence, the case-dependency factor plays a crucial role in shaping the approach to generating synthetic data.

### 3.2. Benefits of using synthetic data

As noted by EDPS, '*synthetic data is gaining traction in ML for training algorithms with labelled data scarcity*'. In other sources, synthetic data emerges as a crucial asset when real-life data is inaccessible or insufficient due to scarcity, lack of variability, or legal constraints such as the GDPR, intellectual property rights or trade secret protection. Synthetic data also assumes a pivotal role in overcoming the labour-intensive and costly nature of labelling real-life data.

<sup>13</sup> AEPD (Spanish DPA), 'Synthetic data and data protection', (November 2023). Available at: <https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-data-and-data-protection>

<sup>14</sup> European Data Protection Supervisor, Tech Champion: Robert Rieman, publication on 'Synthetic Data'. Available at: [https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data\\_en](https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en)

<sup>15</sup> K. El Emam, L. Mosquera, and R. Hoptroff, 'Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data'. O'Reilly Media Inc, (May 2020). Accessible at: [https://cdn.ttgtmedia.com/rms/pdf/Practical\\_Synthetic\\_Data\\_Generation.pdf](https://cdn.ttgtmedia.com/rms/pdf/Practical_Synthetic_Data_Generation.pdf)

<sup>16</sup> AEPD (Spanish DPA), 'Synthetic data and data protection', (November 2023). Available at: <https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-data-and-data-protection>

<sup>17</sup> Please note that this example is for illustration only and that resemblance between real world data and synthetic data is a nuanced subject and its discussion is beyond the scope of this deliverable. See also: Gal, M. S., & Lynskey, O, 'Synthetic Data: Legal Implications of the Data-Generation Revolution', 109 Iowa Law Review, Forthcoming, LSE Legal Studies Working Paper No. 6/2023, (January 2023). Accessible at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4414385](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4414385).

In practical terms, since the data is generated, it can lower the costs and resources involved in collecting the required data. Using ‘dummy’ data for initial AI model training provides developers with a strategic advantage, yielding faster results before transitioning to real data. One of the most promising applications of synthetic data lies in health research and innovation. It is being explored whether virtual, computer-generated patients can prove valuable in the development of medical drugs and devices, potentially providing a way to reduce reliance on human testing and shorten testing times.<sup>18</sup>

Synthetic data may also be employed to address the underrepresentation of certain cases in existing datasets. For instance, recognizing a bias towards predominantly light skin samples in data repositories, a more inclusive set of skin images was created using synthetic data.<sup>19</sup> This initiative aimed to train detection models capable of effectively recognizing potentially malignant skin conditions, such as melanoma, across a spectrum of shades.

### 3.3. Synthetic data in the guidance of Data Protection Authorities (DPAs)

While synthetic data has been researched and used for some time, the guidance on legal requirements for compliance remains limited. Below, we summarize the guidelines issued by the EDPS, AEPD and ICO (UK DPA).

*Table 1 Summary of guidelines issued by the EDPS, AEPD and ICO (UK DPA)*

<p>EDPS Synthetic Data   <a href="https://european-data-protection-supervisor.europa.eu">European Data Protection Supervisor (europa.eu)</a></p> <ul style="list-style-type: none"> <li>• Synthetic data is artificially generated from original data and a model trained to reproduce characteristics and structure of that data.</li> <li>• Privacy assurance assessments should be performed to ensure that the resulting synthetic data is not actual personal data.</li> <li>• This privacy assurance evaluates the extent to which data subjects can be identified in the synthetic data and how much new data about those data subjects would be revealed upon successful identification.</li> <li>• Foreseen impacts:             <ul style="list-style-type: none"> <li>• Positive: enhanced privacy (upon a privacy assurance assessment) and fairness through mitigating bias (synthetic datasets can be manipulated to have a better representativeness of the world).</li> <li>• Negative: complex output control, difficulty in mapping outliers, dependence on original data quality and biases meaning that synthetic data may reflect the biases in original data</li> </ul> </li> </ul>
<p>AEPD (Agencia Espanola Proteccion Datos) <a href="#">Synthetic data and data protection   AEPD</a></p> <ul style="list-style-type: none"> <li>• Synthetic data could be a privacy enhancing technique, when it is used to create non-personal data sets with the same utility than the personal ones.</li> <li>• Should not contain identifiable information even when it is generated from real personal data.</li> <li>• Assessment of a possible risk of re-identification (‘anonymity assessment’) from the created synthetic data set is necessary. Additional privacy enhancing techniques (PETs) such as differential privacy may be needed in some cases.</li> <li>• The creation of synthetic data from real personal data would itself be a processing activity under the GDPR.</li> </ul>

<sup>18</sup> *Ibidem.*

<sup>19</sup> Timo Kohlberger & Yuan Liu, ‘Generating Diverse Synthetic Medical Image Data for Training Machine Learning Models’, (February 2020). Accessible at: <https://blog.research.google/2020/02/generating-diverse-synthetic-medical.html?m=1>

- Synthetic data is not always the right choice: usability must be assessed on a case-by-case basis.

ICO ([chapter-5-anonymisation-pets.pdf \(ico.org.uk\)](#))

- It helps train AI models when access to large datasets is limited; and can be used to generate non-personal data in situations where you cannot share personal data
- If you generate synthetic data from personal data, any inherent biases in the data will be carried through
- It helps comply with data minimisation principle by reducing or eliminating the processing of personal data.
- The more that the synthetic data mimics real data, the greater the utility it has. At the same time, it may be more likely to reveal individuals' personal data.
  - Assessing re-identification risk involved with synthetic data is an ongoing area of development.
- Risks include accuracy, re-identification and vulnerability to attacks.
  - Further additional measures (e.g. differential privacy) may be required to protect against singling out.

### 3.4. Applicability of the GDPR to synthetic data

As follows from the above, the authorities and most researchers agree that synthetic data is not automatically 'private'<sup>20</sup> or placed outside of the realm of data protection laws. Legal considerations predominantly arise when creating synthetic data from real-life datasets containing personal data, as seen for example in medical datasets. In such cases, the process begins with collecting and preparing actual personal data for training AI models that generate synthetic data. From a GDPR perspective, creating synthetic data based on personal data requires processing of the latter.<sup>21</sup> This imposes several requirements. For example, the GDPR principle of data minimization (Article 5.1c) should be implemented by pseudonymizing the input data and removing direct identifiers from it. Another crucial principle is ensuring the integrity and confidentiality of input personal data (Article 5.1f), particularly by incorporating technical and organizational security measures (Article 32) to safeguard it from unlawful disclosure.

As with any personal data processing, there is a need for a legal basis for using input personal data for synthetic data generation. Opinion 05/2014 of the Article 29 Working Party on Anonymisation Techniques<sup>22</sup> states that anonymisation as an instance of further processing of personal data can be compatible with the original purposes of the processing if the result is truly anonymous data. According to some authors, similar argument can be made for synthetic data generation '*provided that the data synthesis is carried out adequately and synthetic data is reliably produced*'<sup>23</sup> or, with a higher standard, that the synthetic data is anonymous (non-personal).

<sup>20</sup> Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S. N., & Weller, 'Synthetic Data - what, why and how?' (May 2022). Accessible at: [https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/Synthetic\\_Data\\_Survey-24.pdf](https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/Synthetic_Data_Survey-24.pdf)

<sup>21</sup> Ganev, Georgi, 'When Synthetic Data Met Regulation', arXiv preprint arXiv:2307.00359v1, (July 2023). Accessible at: <https://arxiv.org/pdf/2307.00359.pdf>

<sup>22</sup> Article 29 Working Party, Opinion 5/2014 on Anonymization Techniques. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

<sup>23</sup> López, C. A. F, 'On synthetic data: a brief introduction for data protection law dummies', European Law Blog, (September 2022). Accessible at: <https://europeanlawblog.eu/2022/09/22/on-synthetic-data-a-brief-introduction-for-data-protection-law-dummies/>

This leads to the imminent question of whether synthetic data is ‘personal data’ governed by the data protection law. On the face of it, one may argue that since the data is purposely disrupted and changed (there is no one-to-one mapping from synthetic records back to the person), it is automatically non-personal. However, there have been studies<sup>24</sup> that indicate that not in all cases sufficient level of anonymization is achieved. Even if the generation of the data was performed on initially de-identified data (where direct identifiers, such as names were removed), there remains a risk that an individual can be indirectly identifiable either from the synthetic data itself or with other available sources<sup>25</sup>

The potential risk becomes especially relevant in cases where a model is vulnerable to ‘overfitting’. In such instances, the model excessively focuses on the details of the training data, essentially memorizing examples from that data and reproducing them in synthetic data<sup>26</sup>. Consequently, this phenomenon exposes a vulnerability in synthetic data, as it has ‘*the capacity to leak information about the data it was derived from*’<sup>27</sup> rendering it susceptible to privacy attacks.

As a result, conducting a thorough assessment of any synthetic data becomes imperative to ascertain its personal or non-personal status. Notably, EDPS emphasized that this assessment should evaluate the extent to which data subjects can be identified in the synthetic data and the amount of new data about those subjects that would be revealed upon successful identification.<sup>28</sup>

Nevertheless, such an assessment is not a straightforward process. From a legal perspective, the assessment of synthetic data under the GDPR is influenced by the ongoing debate on the limits of ‘personal data’. This topic is very complex (refer to recent rulings of CJEU Case C-319/22 and GC T-557/20<sup>29</sup>), resulting in a lack of agreed standards and a potentially expansive definition of ‘personal data.’ Essentially, debates concerning the risk of identification within the GDPR definition of personal data often centre on determining whose perspective should decide if a piece of information qualifies as personal. Additionally, there is a need to establish a threshold for ‘reasonable likenesses’ as a measure to assess the risk of re-identification. Another persistent issue associated with synthetic data involves the potential deduction of sensitive information about an individual, even in cases where the identifiability test fails to yield a positive outcome.

Even if the synthetic data falls short of the anonymity threshold, replacing collected personal data with artificially generated data offers an additional layer of security to personal data. The AEPD and ICO consider synthetic data as a PET which aims to weaken or break the connection between an individual in the original personal data. They propose combining synthetic data with other PETs, such as differential privacy, to enhance privacy protection while retaining utility.

<sup>24</sup> Theresa Stadler, Bristena Oprisanu, Carmela Troncoso, ‘Synthetic Data -- Anonymisation Groundhog Day’, (November 2020). Accessible at: <https://arxiv.org/abs/2011.07018>

<sup>25</sup> Colin Mitchell and Elizabeth Redrup Hill, ‘Are synthetic health data ‘personal data’?’. Accessible at: <https://www.phgfoundation.org/report/are-synthetic-health-data-personal-data#:~:text=We%20found%20that%20regulators%20and,been%20reduced%20to%20remote%20levels>

<sup>26</sup> Ganey, Georgi, ‘When Synthetic Data Met Regulation’, arXiv preprint arXiv:2307.00359v1, (July 2023). Accessible at: <https://arxiv.org/pdf/2307.00359.pdf>

<sup>27</sup> Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S. N., & Weller, ‘Synthetic Data - what, why and how?’ (May 2022). Accessible at: <https://royalsocietypublishing.org/doi/10.1098/rsos.220304>

<sup>28</sup> European Data Protection Supervisor, Tech Champion: Robert Rieman, publication on ‘Synthetic Data’. Accessible at: [https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data\\_en](https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en)

<sup>29</sup> Alexandre Lodie, European Law Blog, ‘Are personal data always personal? Case T-557/20 SRB v. EDPS or when the qualification of data depends on who holds them’, (November 2023). Accessible at: <https://europeanlawblog.eu/2023/11/07/are-personal-data-always-personal-case-t-557-20-srb-v-edps-or-when-the-qualification-of-data-depends-on-who-holds-them/#more-9476>

The qualification of synthetic data has only been briefly mentioned in EU court cases, notably in a CJEU case from 5 December 2023 (Case C-683/21). The case concerned an IT company (ITSS) which developed a Covid tracing app for the Lithuanian Ministry of Health (NVSC). Following the release of the app to the public, NVSC and ITSS were fined as joint controllers for GDPR infringement. However, since the contract between NVSC and ITSS was never finalized there was a dispute on who is the controller of the data collected by the app. One of the circumstances considered was that the ITSS received copies of the personal data collected by the mobile app and used fictitious data ‘with the exception of the telephone numbers of that company’s employees’ for IT testing purposes. In the context on answering the question on what constitutes ‘processing’ CJEU ruled that:

*‘In the light of the foregoing, the answer to the fourth question is that Article 4(2) of the GDPR must be interpreted as meaning that the use of personal data for the purposes of the IT testing of a mobile application constitutes ‘processing’, within the meaning of that provision, unless such data have been rendered anonymous in such a manner that the subject of those data is not or is no longer identifiable, or unless it involves fictitious data which do not relate to an existing natural person.’*

The judgement does not indicate any further details of the ‘fictitious data’ which was used and there are no specific guidelines how to assess if the fictitious data relates to an existing natural person. Still, the verdict is in line with the current stance of certain data protection authorities and privacy researchers.

Synthetic data must be evaluated within the framework of the GDPR, and the privacy implications of any synthetic dataset are highly contingent on the specific context. This perspective is seen as a potential barrier to advancing the use of synthetic data in research. Concerns have been raised regarding the complex legal requirements and the GDPR compliance processes that must be adhered to, which could impede technological progress and hinder the widespread adoption of synthetic data.

The origin of synthetic data is an important factor in assessing whether it qualifies as personal data. When synthetic data is created from original personal data, a crucial trade-off emerges where utility and anonymity are inherently interconnected. The more utility a synthetic dataset provides, the lower its anonymity (meaning the higher the risk of reidentification), and vice versa<sup>30</sup>. Therefore, striking a balance between absolute anonymity and utility preservation is a nuanced task when synthetic data is generated from real personal data, and it is unlikely that a unanimous consensus will emerge asserting that synthetic data is unequivocally non-personal in all instances. Conversely, synthetic data generated through assumptions, bypassing the direct processing of personal data, will not need to face these challenges.

Experts advise that context and practice will have a major influence on the risk of re-identification and argue that DPAs and the community should arrive at ‘appropriate standards and approaches to assessing identifiability of specific synthetic data generation methods, utilizing quantitative metrics as far as possible’<sup>31</sup>

<sup>30</sup> Khaled El Emam, ‘Precaution, ethics and risk: Perspectives on regulating non-identifiable data’, IAPP, (May 2022). Accessible at: <https://iapp.org/news/a/precaution-ethics-and-risk-perspectives-on-regulating-non-identifiable-data/> and López, Cesar Augusto Fontanillo, ‘On the legal nature of synthetic data’, NeurIPS 2022 Workshop on Synthetic Data for Empowering ML Research, (2022). Accessible at: <https://openreview.net/pdf?id=M0KMbGL2yr>

<sup>31</sup> Colin Mitchell and Elizabeth Redrup Hill, ‘Are synthetic health data ‘personal data?’’. Accessible at: <https://www.phgfoundation.org/report/are-synthetic-health-data-personal-data#:~:text=We%20found%20that%20regulators%20and,been%20reduced%20to%20remote%20levels.>

### 3.5. Resulting requirements for FLUTE

It is important to acknowledge that since synthetic data is relatively new, the rules of its use and legal implications in various domains are still in development. Extra caution is advised in scenarios where the data is to be used for training and validation of AI models intended to be classified as medical devices.

The risks associated to synthetic data include the potential for inaccuracies arising from flawed input data or background information<sup>32</sup>, as well as the risk of bias in data creation due to inadequately balanced input information. Additionally, concerns arise about users' ability to understand the underlying logic applied by machine learning in generating synthetic values, raising questions about the transparency and trustworthiness of the data. Ensuring compliance and responsible data management calls for careful consideration.

In the context of the project, in order to be mindful of compliance with the GDPR, the FLUTE partners should incorporate the following good practices related to synthetic data generated from personal data and use in the project:

- Training synthetic data generators:
  - De-identify and minimize the personal data used for training
  - Make sure there is a legal basis for using personal data for training
  - Fulfil transparency obligations towards data subjects (if applicable)
- Security and privacy of synthetic data:
  - Apply security measures to prevent unauthorized access to and disclosure of synthetic data
  - Before further processing (sharing, making public etc.) synthetic data conduct a 'privacy assurance assessment' – make sure that no personal data is leaked
- Quality and bias:
  - Beware of potential bias in training data
  - Decide (discuss) how to verify quality of the synthetic data
- Use of synthetic data
  - Document synthetic data as a source of AI training data
  - Comply with applicable domain specific restrictions (e.g., medical regulations, AI Act).

---

<sup>32</sup> Theresa Stadler, Bristena Oprisanu & Carmela Troncoso, 'Synthetic Data – Anonymisation Groundhog Day' (unpublished manuscript, January 2022). Accessible at: <https://arxiv.org/pdf/2011.07018.pdf>.

## 4. Appropriate safeguards under art. 89 GDPR

### 4.1. Overview

Art. 89(1) GDPR mandates that the processing of personal data for scientific purposes shall be subject to appropriate safeguards, in accordance with the Regulation, for the rights and freedoms of the data subjects. Those safeguards shall ensure that technical and organisational measures are in place to ensure particularly respect for the principle of data minimisation.

GDPR does not define the term ‘appropriate safeguards’ and only provides examples of the safeguards<sup>33</sup>, namely pseudonymisation and anonymisation, failing to give indication as to which other measures may meet this requirement<sup>34</sup>. The provision of art. 89(1) GDPR has not been examined by the CJEU. Regulatory guidance on appropriate safeguards under Art. 89(1) is also scarce, despite the term being present in several opinions of EDPS and EDPB. In a document issued in 2021<sup>35</sup>, EDPB acknowledged the complexity and importance of explaining the meaning of appropriate safeguards and referred to future guidance for an in-depth analysis. This guidance is still to be released.

Based on the wording of the provision, the safeguards to be considered appropriate must: (i) protect rights and freedoms of the data subjects, (ii) lead to implementation of technical and organizational measures which in turn ensure at least the principle of data minimization (including necessity and proportionality), and possibly other art. 5 GDPR principles such as data security and data protection by design and by default (iii) be compliant with GDPR.

### 4.2. Resulting requirements for FLUTE

FLUTE Platform should investigate the requirements for appropriate safeguards and implement measures which have been indicated as such safeguards. A non-exhaustive list of such safeguards includes:

- Anonymization and/or pseudonymization of data to be made available for research via in the FLUTE Platform,
- Technical measures which aim at minimising the amount of data processed, including, for example federated learning (FL), PETs and privacy budgets,
- Security features of the Platform and the local nodes which prevent personal data breaches,
- Assurances regarding legal and ethical compliance of conducted research activities, including authorization and authentication of the users of the Platform and their research goals.

As indicated in Section 2.4 all of the safeguards indicated above were included in the prototype of FLUTE platform.

---

<sup>33</sup> In fact, the term “appropriate safeguards” appears in the GDPR 37 times and has different meaning depending of the context. For instance, in Art. 6(4) GDPR “appropriate safeguards” are listed as one of the factors to be considered when evaluating compatibility for further processing of data. In Art. 40, the GDPR refers to codes of conduct as a type of “appropriate safeguards” within the framework of personal data transfers to third countries or international organisations. Art. 46 mentions about “appropriate safeguards” in the context of transfers of data to non-adequate countries.

<sup>34</sup> Ciara Staunton and others, ‘Appropriate Safeguards and Art. 89 of the GDPR: Considerations for Biobank, Databank and Genetic Research’ (2022) 13 *Frontiers in Genetics*

<sup>35</sup> EDPB, Response to the Request from the European Commission for Clarifications on the Consistent Application of the GDPR, Focusing on Health Research, adopted on 2 February 2021. In another part of this document, however, EDPB briefly mentioned that informed consent for participation in the medical research project could be perceived as an additional safeguard.

## 5. Regulation on Artificial Intelligence (AI)

### 5.1. Responsible and trustworthy AI

Trustworthy AI has 3 main components, which should be met throughout the entire lifecycle of the system:

- it should be **lawful**, complying with all applicable laws and regulations;
- it should be **ethical**, ensuring adherence to ethical principles and values; and
- it should be **robust**, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.<sup>36</sup>

It is advisable that FLUTE follows the ethics guidelines for trustworthy AI presented by the High-Level Expert Group on Artificial Intelligence<sup>37</sup>, which puts forward 7 requirements for trustworthy AI:

1. Human agency and oversight: AI systems should support human autonomy and decision-making, as prescribed by the principle of respect for human autonomy. AI systems should support the user's agency, foster fundamental rights and allow for human oversight.
2. Technical robustness and safety: AI systems should be developed with a preventive approach to risks and minimising or preventing harm. They should be secure and resilient to attacks, should have safeguards that enable a fallback plan in case of problems and should produce accurate, reproducible and reliable results.
3. Privacy and data governance: AI systems should guarantee privacy and data protection throughout a system's entire lifecycle. Adequate privacy protection also necessitates data governance that covers the quality and integrity of the data used as well as data access protocols.
4. Transparency: AI systems should be transparent, in a way that data sets and processes are traceable and explainable. Moreover, people need to be informed when they are interacting with an AI system.
5. Diversity, non-discrimination and fairness: AI systems should ensure inclusion and diversity throughout the entire system's life cycle. It should avoid unfair biases which could lead to discrimination and should be designed in an accessible and universal way.
6. Environmental and societal well-being: AI systems should be sustainable and ecologically responsible to the greatest extent possible. They should be used to benefit all human beings, including future generations.
7. Accountability: AI systems should be accountable. This necessitates that mechanisms be put in place to ensure responsibility and accountability for AI systems and their outcomes, both before and after their development, deployment and use.

The general principles of ALTAI have been developed into more concrete checklists. Some of the sources of such checklists include:

<sup>36</sup> High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI (2019). Available at: <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>

<sup>37</sup> Ibid. See also High-Level Expert Group on Artificial Intelligence, Assessment List for Trustworthy AI (ALTAI) (2020). Available at: <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>

- AI toolkit designed by the ICO as practical support in reducing the risks to individuals' rights and freedoms caused by their own AI systems<sup>38</sup>.
- FUTURE-AI is an international, multi-stakeholder initiative for defining and maintaining concrete guidelines that will facilitate the design, development, validation and deployment of trustworthy AI solutions in medicine and healthcare based on six guiding principles: Fairness, Universality, Traceability, Usability, Robustness and Explainability. Future-AI proposed an assessment checklist consisting of concrete and actionable questions that guide developers, evaluators and other stakeholders in delivering medical AI tools that are trustworthy and optimised for real-world practice<sup>39</sup>.
- CNIL (French DPA) has published a self-assessment guide for artificial intelligence (AI) systems intended for controllers and processors implementing personal data processing based on AI systems<sup>40</sup>.

## 5.2. Overview of the EU Artificial Intelligence Act (AIA)

After three years of legislative works, the AIA<sup>41</sup> was finally published on 12 July 2024. It is the first-ever legal framework on AI, which lays down harmonised rules for developers and users of certain AI systems. AI system under the AIA refers to a machine-based system that:

- is designed to operate with varying levels of autonomy and that
- may exhibit adaptiveness after deployment, and that
- for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

Key characteristic of AI systems is their capability to infer. AI systems can operate independently without human intervention, by self-learning and adapting over time based on new data. However, the term does not cover systems that are based on the rules defined solely by natural persons to automatically execute operations. In February 2025, the EC published guidelines on AI system definition to facilitate the first AIA's rules application<sup>42</sup>. These guidelines provide further analysis to the abovementioned elements. Considering the wide variety of AI systems, the guidelines are not exhaustive in line with recital 12 of the AIA, which clarifies that the notion of an 'AI system' should be clearly defined while providing 'the flexibility to accommodate the rapid technological developments in this field'. Moreover, the definition of the AI system adopts a lifecycle-based perspective encompassing two main phases: the pre-deployment or 'building' phase of the system and the post deployment or 'use' phase of the system. Specific elements of the definition may appear at one phase, but it is possible that they do not persist across both phases.

<sup>38</sup> ICO, AI and data protection risk toolkit. Available at: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/ai-and-data-protection-risk-toolkit/>

<sup>39</sup> <https://future-ai.eu/checklist/>.

<sup>40</sup> CNIL, Self-assessment guide for artificial intelligence (AI) systems, Available at: <https://www.cnil.fr/en/self-assessment-guide-artificial-intelligence-ai-systems>

<sup>41</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

<sup>42</sup> European Commission, Guidelines on AI system definition established by Regulation (EU) 2024/1689 (AI Act). Available at: <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application#:~:text=By%20issuing%20guidelines%20on%20the%20AI%20system%20definition%2C,on%20the%20AI%20system%20definition%20are%20not%20binding>

### 5.3. Limitations of AIA and risk-based approach

The AIA applies to AI systems put on the EU market. This means that also developers located outside of the EU will have to comply with its rules, if they want to make their AI products available in the EU. There are however some **exceptions**. For example, AIA does not apply to:

- AI systems exclusively for military, defence or national security purposes,
- AI systems released under free and open-source licences, unless they are placed on the market or put into service as high-risk AI systems (HRAIs), if they fall in the scope of prohibited AI practices or of specific transparency obligations for providers and deployers of certain AI systems,
- users who are natural persons using AI systems for their purely personal non-professional activity.

Moreover, the AIA applies the so-called ‘risk-based approach’, which means that the higher the risk associated with an AI system, the more rigorous the risk management and compliance obligations. Thus, different provisions of regulation apply to AI systems depending on their risk.

The table below provides types of AI systems, based on the categories of risks, presents examples of AI systems in each category, and gives a high-level overview of the main legal requirements applicable to them.

Table 2 Risk Based approach in the AIA

Type	Examples	Main requirements
<b>Prohibited AI systems:</b>  <b>Unacceptable risk as it violates EU fundamental rights and values.</b>	For example, AI systems that: <ul style="list-style-type: none"> <li>• deploy subliminal techniques beyond a person’s consciousness or purposefully manipulative or deceptive techniques,</li> <li>• exploit the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation,</li> <li>• perform ‘social scoring’ - evaluation or classification of natural persons based on their social behaviour or known, inferred or predicted personal or personality characteristics</li> <li>• predict the risk of a natural person committing a criminal offence, based on profiling or assessing personality traits and characteristics.</li> </ul>	Deployment of such AI systems in the EU is forbidden.
<b>High-risk AI systems:</b>  <b>High risk as it impacts health, safety or fundamental rights</b>	<ul style="list-style-type: none"> <li>• Stand alone AI systems (Annex III)</li> <li>• AI systems which are safety components of products, or products themselves (covered by the Union harmonisation legislation listed in Annex I) – examples: diagnostics systems and systems supporting human decisions</li> </ul>	Strict requirements for providers and of deployers (users). See below for details.

	See below for details.	
<b>Sui generis risk AI systems:</b> <b>Risk of deception</b>	Chatbots, deepfakes	Mostly transparency requirements to avoid risk of deception of the public. See below for details.
<b>Other AI systems:</b> <b>Minimal risk</b>	Spam filters, document editors, recommender systems, etc.	Not regulated by the AIA, but consumer protection and product safety rules apply.

Additionally, the AIA regulates general-purpose AI models. Those are models that have a wide range of possible uses (for example, large language models such as GPT-4, etc.). They are subject to a tiered approach depending on whether it is a general purpose AI model with systemic risk or a ‘normal’ GPAI model.

### 5.4. High-Risk AI Systems (HRAIS)

HRAIS are the most heavily regulated by the AIA. The table below explains the differences between two types of HRAIS provided in the AIA.

Table 3 HRAIS

	Stand alone AI systems (Annex III)	AI systems which are safety components of products, or products themselves (Annex I)
<b>Conditions</b>	<p>Annex III identifies 8 areas where the use of AI is particularly sensitive and provides specific use cases for each. An AI system is classified as HRAIS if it is intended for one of these use cases.</p> <p>Examples include:</p> <ul style="list-style-type: none"> <li>a) AI systems used for access to essential public and private services and benefits (e.g., healthcare), creditworthiness assessments of individuals, and risk evaluation and pricing in life and health insurance.</li> <li>b) AI systems employed for biometric identification, biometric categorization, and emotion recognition, except where prohibited.</li> </ul> <p><b>Note:</b> The list may be amended by European Commission’s (EC) delegated acts.</p>	<p>Both conditions must be fulfilled:</p> <ul style="list-style-type: none"> <li>a) AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation listed in Annex I <u>and</u></li> <li>b) AI system or product is required to undergo a third-party conformity assessment</li> </ul> <p><b>Note:</b> Medical Device Regulation (MDR) as well as In Vitro Medical Device Regulation (IVDR) are included in Annex I.</p>

<p><b>Notes and exceptions</b></p>	<p><b>Exceptions:</b></p> <ul style="list-style-type: none"> <li>• where the AI system does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making;</li> <li>• performs a narrow procedural task</li> </ul> <p>The reliance on this exception must be documented.</p> <p><b>Example:</b> AI system that transforms unstructured data into structured data.</p> <p><b>Note:</b> AI system will automatically be considered to be a HRAI if the AI system performs profiling of individuals.</p>	<p><b>Note:</b> Where a HRAI that is a safety component of a product which falls within the scope of Union harmonisation legislation based on the New Legislative framework<sup>43</sup> is not placed on the market or put into service independently from the product, the <b>product manufacturer</b> defined in that legislation should comply with the obligations of the provider established in the AIA and should, in particular, ensure that the AI system embedded in the final product complies with the requirements of the AIA.</p>
<p><b>Examples of HRAIs from the health domain</b></p>	<p>AI systems intended to:</p> <ul style="list-style-type: none"> <li>• be used as emergency healthcare patient triage systems;</li> <li>• be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services.</li> </ul>	<ul style="list-style-type: none"> <li>• Medical devices in classes IIa, IIb, III; some MD in class I, if placed on the market in sterile condition, having a measuring function or being reusable surgical instruments;</li> <li>• In vitro diagnostic medical devices requiring third party conformity assessment.</li> </ul>

### 5.5. Requirements for HRAIS

As for the main actors in the AIA, these are the ‘provider’ and the ‘deployer’. The ‘provider’ refers to the entity that develops and places or puts into service an AI system on the EU market and is either established within the EU or a third country, where the output produced by the AI system is used in the EU. The ‘deployer’ refers to the entity that uses an AI system in the EU or a third country under its authority, and it is established within the EU or a third country, where the output produced by the AI system is used in the EU.

Requirements that apply to HRAIS regard, among others, risk management, the quality and relevance of data sets used, technical documentation and record-keeping, transparency and the provision of information to deployers, human oversight, and robustness, accuracy and cybersecurity.

<sup>43</sup> [https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework\\_en](https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework_en)

Table 4 Requirements for HRAIs

Requirement	Conditions
Risk management system	<p>Continuous iterative process subject to systematic review and updating run throughout the lifecycle of the HRAI aiming at the:</p> <ul style="list-style-type: none"> <li>• Identification and analysis of the known and the reasonably foreseeable risks;</li> <li>• Estimation and evaluation of risks under conditions of reasonably foreseeable misuse;</li> <li>• Evaluation of other risks possibly arising, on the basis of post-market monitoring analysis;</li> <li>• Adoption of appropriate and targeted risk management measures;</li> </ul> <p>HRAIs shall be tested for the purpose of identifying the most appropriate and targeted risk management measures.</p>
Data and data governance	<p>HRAIs involving the training of AI models with data shall be developed on the basis of training, validation and testing data sets that meet specific quality criteria, such as representativeness and accuracy.</p> <p>Data governance and management practices must concern in particular:</p> <ul style="list-style-type: none"> <li>• the relevant design choices;</li> <li>• data collection processes and the origin of data, and in the case of personal data, the original purpose of the data collection;</li> <li>• relevant data-preparation processing operations, such as annotation, labelling, cleaning, updating, enrichment and aggregation;</li> <li>• the formulation of assumptions, in particular with respect to the information that the data are supposed to measure and represent;</li> <li>• an assessment of the availability, quantity and suitability of the data sets that are needed;</li> <li>• examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations;</li> <li>• appropriate measures to detect, prevent and mitigate possible biases identified according to point (f);</li> <li>• the identification of relevant data gaps or shortcomings that prevent compliance with this Regulation, and how those gaps and shortcomings can be addressed.</li> </ul> <p>The right to privacy and to protection of personal data must be guaranteed throughout the entire lifecycle of the AI system. In this regard, the principles of data minimisation and data protection by design and by default, as set out in GDPR, are applicable when personal data are processed. Measures taken by providers to ensure compliance with those principles may include not only anonymisation and encryption, but also the use of technology that permits algorithms to be brought to the data and allows training of AI systems without</p>

	<p>the transmission between parties or copying of the raw or structured data themselves<sup>44</sup>. The latter recommendation pertains to, for example, FL.</p>
Technical documentation	<p>Technical documentation shall be drawn up before that system is placed on the market or put into service and shall be kept up-to date;</p> <p>It shall contain, at a minimum, the elements set out in Annex IV, .e.g. the system’s intended purpose, instructions for use for the deployer, description of the software and hardware etc;</p> <p>For HRAIs covered by the Union harmonisation legislation of Annex I, a single set of technical documentation shall be drawn up containing all the information mentioned above, and information required under those legal acts.</p>
Record-keeping	<p>HRAIs shall allow for the automatic recording of events (logs) over the lifetime of the AI system;</p> <p>Logging capabilities shall enable the recording of events relevant for aspects such as the identification of situations that may result in the HRAI presenting a risk, or the facilitation of post-market monitoring;</p> <p>For AI systems under Annex III, the logging capabilities shall provide, at a minimum recording of the period of each use of the system, the input data etc.</p>
Transparency and provision of information to deployers	<p>HRAIs shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system’s output and use it appropriately;</p> <p>Manufacturers shall make sure that HRAIs are accompanied by instructions for use containing at least the following information:</p> <ul style="list-style-type: none"> <li>• the identity and the contact details of the provider and, where applicable, of its authorised representative;</li> <li>• the characteristics, capabilities and limitations of performance of the HRAI, e.g., its intended purposes, the level of accuracy, any known or foreseeable circumstance which may lead to risks, specifications for the input data, etc.</li> </ul>
Human oversight	<p>With the aim of preventing or minimising risks, HRAIs shall be designed and developed in such a way, that they can be effectively overseen by natural persons during the period in which they are in use;</p> <p>The oversight measures shall be commensurate with the risks, level of autonomy and context of use of the HRAI and applicable already in the research phase;</p> <p>The HRAI shall be provided to the deployer in such a way that natural persons to whom human oversight is assigned are enabled to:</p> <ul style="list-style-type: none"> <li>• properly understand the relevant capacities and limitations of the HRAI and be able to duly monitor its operation;</li> <li>• remain aware of the possible tendency of automatically relying or over-relying on the output produced by a HRAI;</li> <li>• correctly interpret the HRAI’s output;</li> </ul>

<sup>44</sup> Recital 69 AIA.

	<ul style="list-style-type: none"> <li>decide, in any particular situation, not to use the HRAI or to otherwise disregard, override or reverse its output;</li> <li>intervene in the operation of the HRAI or interrupt the system through a 'stop' button or a similar procedure.</li> </ul>
<p>Accuracy, robustness and cybersecurity</p>	<p>HRAIs shall be designed and developed in such a way that they achieve an appropriate level of accuracy, robustness, and cybersecurity, and that they perform consistently in those respects throughout their lifecycle;</p> <p>The EC shall, in cooperation with relevant stakeholders and organisations such as metrology and benchmarking authorities, encourage, as appropriate, the development of benchmarks and measurement methodologies;</p> <p>The levels of accuracy shall be declared in the accompanying instructions of use;</p> <p>HRAIs shall be as resilient as possible regarding errors, faults or inconsistencies. Their robustness may be achieved through technical redundancy solutions, such as backup or fail-safe plans. For HRAIs that continue to learn after being placed on the market or put into service, their development shall eliminate feedback loops;</p> <p>HRAIs shall be resilient against attempts by unauthorised third parties to alter their use, outputs or performance by exploiting system vulnerabilities.</p>
<p>Quality management system</p>	<p>Providers of HRAIs shall put a quality management system, including aspects, such as a strategy for regulatory compliance, techniques, procedures and systematic actions the design, design control and design verification of the HRAI, the handling of communication with national authorities, reporting procedures, etc.;</p> <p>The implementation of the above aspects shall be proportionate to the size of the provider's organisation.</p>
<p>Documentation-keeping</p>	<p>The provider shall, for a period ending 10 years after the HRAI has been placed on the market or put into service, keep at the disposal of the national competent authorities:</p> <ul style="list-style-type: none"> <li>the technical documentation;</li> <li>the documentation concerning the quality management system;</li> <li>the documentation concerning the changes approved by notified bodies, where applicable;</li> <li>the decisions and other documents issued by the notified bodies, where applicable;</li> <li>the EU declaration of conformity.</li> </ul>
<p>Corrective actions and duty of information</p>	<p>Providers of HRAIs which consider or have reason to consider that a HRAI that they have placed on the market or put into service is not in conformity with this Regulation shall immediately take the necessary corrective actions to bring that system into conformity, to withdraw it, to disable it, or to recall it, as appropriate.</p> <p>They shall inform the distributors of the HRAI concerned and, where applicable, the deployers, the authorised representative and importers accordingly.</p>

Cooperation with competent authorities	Providers of HRAIs shall, upon a reasoned request by a competent authority, provide that authority all the information and documentation necessary to demonstrate the conformity of the HRAI with the requirements for HRAIs, or allow access to the automatically generated logs.
Authorised representatives	<p>Prior to making their HRAIs available on the Union market, providers established in third countries shall, by written mandate, appoint an authorised representative which is established in the Union;</p> <p>The authorised representative shall perform the tasks specified in the mandate received from the provider, such as verifying the EU declaration of conformity, keeping at the disposal of competent authorities the contact details of the provider, etc.</p>
Fundamental rights impact assessment (FRIA)	<p>Obligation for deployers that are bodies governed by public law, or are private entities providing public services for specific stand-alone HRAIs under Annex III (e.g., access to and enjoyment of essential private and essential public services and benefits) prior to the first use of the HRAI;</p> <p>The assessment shall consist of:</p> <ul style="list-style-type: none"> <li>• a description of the deployer’s processes in which the HRAI will be used in line with its intended purpose;</li> <li>• a description of the period of time within which and the frequency with which each HRAI is intended to be used;</li> <li>• the categories of natural persons and groups likely to be affected by its use in the specific context;</li> <li>• the specific risks of harm likely to have an impact on the categories of persons or groups of persons identified pursuant to the point above, taking into account the information given by the provider pursuant to Article 13 of the AIA;</li> <li>• a description of the implementation of human oversight measures according to the instructions for use;</li> <li>• the measures to be taken where those risks materialize, including arrangements for internal governance and complaint mechanisms.</li> </ul> <p>If a DPIA already meets some FRIA obligations, the FRIA will complement the DPIA. Other relevant prior impact assessments can also inform the FRIA;</p> <p>After performing the FRIA, deployers must notify the designated market surveillance authority in each EU Member State of the assessment results.</p>
Conformity assessment	<p>For HRAIs covered by the Union harmonisation legislation listed in Annex I (e.g., Medical Devices Regulation, In Vitro Diagnostic Medical Devices Regulation), the provider shall follow the relevant conformity assessment procedure as required under those legal acts.</p> <p>HRAIs that have already been subject to a conformity assessment procedure shall undergo a new conformity assessment procedure in the event of a substantial modification, regardless of whether the modified system is intended to be further distributed or continues to be used by the current deployer.</p>
EU declaration of conformity	The provider shall draw up a written machine readable, physical or electronically signed EU declaration of conformity for each HRAI, and keep it

	<p>at the disposal of the national competent authorities for 10 years after the HRAI has been placed on the market or put into service;</p> <p>Where HRAIs are subject to other Union harmonisation legislation which also requires an EU declaration of conformity, a single EU declaration of conformity shall be drawn up;</p> <p>By drawing the EU declaration of conformity, the provider assumes responsibility for compliance.</p>
CE marking of conformity	<p>The CE marking shall be affixed visibly, legibly and indelibly for HRAIs, or the packaging or to the accompanying documentation, as appropriate;</p> <p>Where HRAIs are subject to other Union law which also provides for the affixing of the CE marking, the CE marking shall indicate that the HRAI also fulfil the requirements of that other law.</p>
EU database registration	<p>Before placing on the market or putting into service a HRAI listed in Annex III, the provider or, where applicable, the authorised representative shall register themselves and their system in the EU database set up by the EC and Member States;</p> <p>The EU database shall contain personal data only in so far as necessary for collecting and processing information in accordance with this Regulation. That information shall include the names and contact details of natural persons who are responsible for registering the system and have the legal authority to represent the provider or the deployer, as applicable.</p>
Post-market monitoring	<p>Providers shall establish and document a post-market monitoring system in a manner that is proportionate to the nature of the AI technologies and the risks of the HRAI;</p> <p>The post-market monitoring system shall be based on a post-market monitoring plan, which forms part of the technical documentation;</p> <p>For HRAIs covered by the Union harmonisation legislation under Annex I, where a post-market monitoring system and plan are already established under that legislation, providers shall have a choice of integrating the necessary elements under the AIA into systems and plans already existing under that legislation, provided that it achieves an equivalent level of protection.</p>
Reporting of serious incidents	<p>Providers of HRAIs placed on the Union market shall report any serious incident to the market surveillance authorities of the Member States where that incident occurred.</p> <p>The report referred to in paragraph 1 shall be not later than 15 days after the provider or, where applicable, the deployer, becomes aware of the serious incident. In case of a widespread infringement the report shall be provided not later than two days after the provider or, where applicable, the deployer becomes aware of that incident.</p> <p>Following the reporting of a serious incident the provider shall, without delay, perform the necessary investigations in relation to the serious incident and the AI system concerned, including a risk assessment of the incident, and corrective action.</p>

## 5.6. HRAIS and harmonised standards

Standards are a voluntary mechanism to ensure that products and services are interoperable with one another, and safe to use. They address risks to health, safety and fundamental rights of individuals throughout product lifecycle. They also have a key role to play in enabling innovation, as they provide a common framework on which to build by setting out the essential characteristics of a product or service and defining common vocabularies. Standards should define, to the extent possible, horizontal requirements, i.e. requirements that are applicable to various types of AI systems across sectors. These can be complemented, when necessary, with requirements that apply to specific sectors or to specific types of systems. Even though standards are not mandatory, once published in the Official Journal of the European Union (OJEU), they serve as a presumption of conformity for the AI systems that comply with them.

Currently, there is ongoing work for the development of standards for AI systems at international level, through the ISO/IEC SC-42 committee, and at European level, through the CEN/CENELEC JTC-21 committee. In the context of the latter committee, the EC has submitted a request for standardisation deliverables corresponding to the requirements for HRAIs under the AIA<sup>45</sup>. Even though initially planned for spring 2025, the European standards are already expected to be delayed sometime before August 2026 when the obligations for HRAIs become applicable.

## 5.7. Application of the AIA

The AIA was published in the EU's Official Journal on 12 July 2024 and has entered into force on 2 August 2024. As a general rule, AIA will apply from 2 August 2026. However, some notable exceptions apply. For example, the obligations for high-risk AI systems will apply as of 2 August 2027.

## 5.8. AIA and scientific research exceptions

The goal of AIA is to support innovation and respect the freedom of science. Hence, the AIA does not apply to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development. Moreover, the regulation does not apply to any research, testing or development activity regarding AI systems or models prior to their being placed on the market or put into service. Such activities must however be conducted in accordance with applicable Union law. When an AI system is placed on the market as a result of a research and development activity, or following the application of provisions on AI regulatory sandboxes and testing in real world conditions of AIA apply.

In the context of AIA, 'testing in real-world conditions' means the temporary testing of an AI system for its intended purpose in real-world conditions outside a laboratory or otherwise simulated environment, with a view to gathering reliable and robust data and to assessing and verifying the conformity of the AI system with the requirements of the AIA. Testing in real-world conditions does not qualify as placing the AI system on the market or putting it into service within the meaning of the AIA, provided that all the applicable conditions are fulfilled.

## 5.9. Resulting requirements for FLUTE

Based on the summary of AIA provided above, the most relevant points for the FLUTE project are as follows:

---

<sup>45</sup> For more information see: European Commission, JRC Publications Repository - Harmonised Standards for the European AI Act, Available at: [https://publications.jrc.ec.europa.eu/repository/bitstream/JRC139430/JRC139430\\_01.pdf](https://publications.jrc.ec.europa.eu/repository/bitstream/JRC139430/JRC139430_01.pdf)

- It is essential to first assess whether the AI systems developed in the context of FLUTE project fall under the definition of the ‘AI system’ and subsequently identify the appropriate classification category.
- It is important to ensure that the AI models do not copy/extract/modify data from the FLUTE private repositories (local nodes).
- During the project lifetime, the AIA does not apply directly to scientific research, hence the specific obligations imposed thereby will not apply; however, any research and development activity should be carried out in accordance with recognized ethical and professional standards for scientific research and in accordance with all applicable Union law.
- Despite the ‘scientific research exception’ it is likely that many of the AIA’s provisions will still have a strong impact on AI research, given the need to anticipate market placement or test new tools in real-world conditions. From the beginning developers should consider how their AI systems will meet the requirements under the AIA to minimize future modifications. In particular, AI systems that are safety components of products, or which are products themselves, and fall under certain Union harmonisation legislation, should be classified as high-risk under the AIA if the product undergoes a third-party conformity assessment. This includes in particular medical devices and in vitro diagnostic medical devices.
- Considering that AI products which are to be placed on the market, have to meet the obligations of the AIA, implementing those requirements already at a research level, will greatly facilitate the process and make life easier for the companies/AI providers. Those most relevant provisions are in particular:
  - Classification rules for AI Systems (Sections 5.3 and 5.4): researchers will have to assess whether the systems they use or currently develop fall under the definition of an AI system. More specifically, partners working on FLUTE AI tools should make an assessment whether those AI systems, either during or after the end of the project, may be qualified as high-risk AI systems, given their intended use. If this will be the case, their manufacturers will need to ensure compliance with both the MDR/IVDR and the corresponding requirements under the AIA (as the MDR and IVDR will continue to apply to healthcare AI alongside the AIA)
  - While developing the high-risk AI systems, the respective partners should observe the requirements outlined above in Section.
  - While collecting and curating the data used for the training of the AI tools which may become HRAIS in the future, provisions on data and data governance (Section 5.5) should be considered. For instance, the project should focus on preventing or reducing data infringements, erroneous decisions-making and bias through an appropriate data governance approach. This includes having a documentation of personal data usage, and assessments of and techniques to ensure the availability and quality of training and input data. This can help protect privacy, ensure quality of output and prevent erroneous decisions-making including algorithmic bias, which is the major ethical concern regarding AI systems.
- Researchers involved in an AI research project should understand the future requirements to draw up the technical documentation of any potential high risk AI systems in order to be able to gradually collect and record relevant information. This was addressed through a dedicated AIA workshop which was organized in November 2024.

- Partners should closely follow the development of standards for AI systems and examine the possibility of adhering to the most suitable ones to facilitate the regulatory and market acceptance of project outcomes.
- To the extent relevant to the goals of the project, the partners should observe its implementation as well as drafting of the delegated acts and aim to participate in the development of the relevant standards and specifications, in particular related to security of AI.
- Lastly, project partners should be mindful of standardisation opportunities for project outcomes. Even though currently standardisation initiatives for AI at European and international level are at rather initial stages, FLUTE project benefited from the HS Booster, a European Commission initiative that provides expert services to European projects to help them to increase and valorise project results by contributing to the creation or revision of standards. Technical partners, supported by TLX, participated in a series of calls with a dedicated advisor to understand which aspects of the project results are eligible for standardisation, under which conditions, and the procedures to be followed. Project partners should assess the relevant information and decide on the next steps. The project standardization partner, HL7, is also exploring ways to standardize AI access to data, for both research and care.

## 6. Cybersecurity laws

### 6.1. NIS2

#### 6.1.1. Overview

The NIS2 Directive<sup>46</sup> was published on 27 December 2022 and the Member States were obligated to publish the measures necessary to comply with this Directive by 17 October 2024 (this process is however still ongoing).

NIS2 replaces the previous directive on security of network and information systems (the NIS Directive) and sets the baseline for cybersecurity risk management measures and reporting obligations across all sectors that are covered by the directive, such as energy, transport, health and digital infrastructure. To achieve this, it sets out minimum rules for a regulatory framework and lays down mechanisms for effective cooperation among relevant authorities in each Member State.

NIS2 updates the list of sectors that are within the directive's scope – including healthcare providers, IT-managed service providers, courier services, manufacturers, waste and water management providers, public administration entities, digital infrastructure providers, social networks and electronic communication service providers. The covered entities will be divided to categories of 'essential' entities (outlined in Annex I of NIS2) and 'important' entities (outlined in Annex II of NIS2).

#### 6.1.2. Scope of NIS2

Organisations subject to the NIS2 regime will be obliged to 'take appropriate and proportionate technical, operational and organisational measures to manage the risks posed to the security of network and information systems which those entities use for their operations or for the provision of their services, and to prevent or minimise the impact of incidents on recipients of their services and on other services'. Those measures include, in particular: policies on risk analysis and information system security, policies on incident handling, access control policies, use of multi-factors authentication or continuous authentication solutions and supply chain security. Importantly, cyber risk management measures include supply chain diligence (supply chain security, including security-related aspects concerning the relationships between each entity and its direct suppliers or service providers). This practically expands the impact of NIS2 beyond the directly regulated 'essential and important entities' to their suppliers and service providers. Furthermore, it mandates personal liability of the management of the organization for failure to comply with its cybersecurity obligations under the legislation. Lastly, it enhances the requirements on reporting cyber-incidents to national computer security incident response teams (CSIRTs) or regulators.

#### 6.1.3. Resulting requirements for FLUTE

NIS2 applies to organizations, rather than products or services, hence it will not be directly applicable to any results of the project, such as the FLUTE Platform or AI models.

---

<sup>46</sup> Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)

## 6.2. Cyber Resilience Act

### 6.2.1. Overview

The EU Cyber Resilience Act (CRA)<sup>47</sup> establishes mandatory horizontal cybersecurity requirements for hardware and software products, introducing common standards applicable throughout a device's lifecycle. It was published on 20 November 2024 and entered into force on 10 December 2024. It will apply as of 11 December 2027 with the exception of manufacturers' reporting obligations that will apply of 11 September 2026.

The CRA aims to ensure that:

- products with digital elements placed on the EU market will have fewer vulnerabilities and that manufacturers remain responsible for cybersecurity throughout a product's lifecycle;
- improve transparency on security of hardware and software products that allows users to make more informed choices;
- business users and consumers benefit from better protection;
- there is a harmonised EU cybersecurity framework for essential cybersecurity requirements, consistent with other EU laws, such as the NIS2 Directive and the GDPR.

Before the CRA, there were no overarching EU-level cybersecurity requirements for all digital devices; existing regulations, such as the EU Cybersecurity Act, applied only to specific products or sectors. To bridge this gap, the CRA applies to products with digital elements made available on the EU market, the intended purpose or reasonably foreseeable use of which includes a direct or indirect logical or physical data connection to a device or network. CRA provides for uniform cybersecurity requirements while tailoring conformity assessment methods based on risk levels. It primarily targets manufacturers, imposing cybersecurity obligations on the planning, design, development, production, delivery and maintenance phase of digital devices. Non-compliant devices will be barred from entering the EU market.

### 6.2.2. Scope

The CRA covers 'products with digital elements' referring to 'software or hardware products and their remote data processing solutions, including software or hardware components being placed on the market separately'. Hardware products may include smartphones, laptops and smartwatches, while software products may include, for example, mobile gaming apps.

However, there are several exceptions, providing that the CRA is not applicable to such products:

- Services that are not linked to a specific product: For example, software as a service (SaaS)<sup>48</sup>, platform as a service (PaaS)<sup>49</sup>, and infrastructure as a service (IaaS)<sup>50</sup> are excluded from the scope of the CRA.

---

<sup>47</sup> Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act) (Text with EEA relevance)

<sup>48</sup> SaaS is a complete software solution provided by a cloud service provider on a pay-as-you-go basis that allows users to connect to and use cloud-based apps over the Internet (e.g., email, calendaring tools). For definition see What is SaaS? Software as a Service | Microsoft Azure

<sup>49</sup> PaaS is provided by a cloud service provider as a complete development and deployment environment in the cloud, with resources that enable the delivery of simple cloud-based apps up to sophisticated, cloud-enabled enterprise applications. For the definition see What is PaaS? Platform as a Service | Microsoft Azure

<sup>50</sup> IaaS is a type of cloud computing service that offers essential compute, storage, and networking resources on demand, on a pay-as-you-go basis. For the definition see What is IaaS? Infrastructure as a Service | Microsoft Azure

Nonetheless, when these services are necessary for the product with digital elements to fulfil its functions, then they fall in the ambit of the CRA;

- Non-commercial free software and open-source software: It refers to software whose source code is openly shared, and its licensing provides for all rights to make it freely accessible, usable, modifiable and redistributable. Only when such software is made available on the market, and therefore supplied for distribution or use in the course of a commercial activity, it should fall within the scope of the CRA;
- Products covered by specific EU harmonisation legislation, such as the MDR.

Other exceptions include, for example, spare parts, prototypes, unfinished software made available for testing or products developed or modified exclusively for national security or defence purposes or specifically designed to process classified information.

### 6.2.3. Essential requirements

Once the CRA is applicable, to be available on the market, products with digital elements will have to meet essential cybersecurity requirements specified in Annex I (e.g., have a secure by default configuration, comply with the GDPR principle of data minimisation, reduce the impact of an incident by using appropriate exploitation mitigation mechanisms, etc.), and be properly installed, maintained, used for their intended purpose, and, where applicable, the necessary security updates must be installed.

Products with digital elements that are certified or for which a statement of conformity has been issued under a European cybersecurity scheme pursuant to Regulation (EU) 2019/881 (EU Cybersecurity Act), shall be presumed to be in compliance with the essential cybersecurity requirements of the CRA in so far as the European cybersecurity certificate or statement of conformity or parts thereof cover those requirements.

### 6.2.4. Obligations of manufacturers, importers, and distributors<sup>51</sup>

The CRA applies to manufacturers, importers, and distributors placing products with digital elements on the EU market, even when they are based outside the EU.

Manufacturers are primarily targeted by the CRA. Their main obligations include:

- Cybersecurity risk assessment: Manufacturers will have to perform a cybersecurity risk assessment, as part of the technical documentation, and take its outcome into account during the planning, design, development, production, delivery and maintenance phases of the product;
- Vulnerability handling requirements: Manufacturers will have to establish and regularly update appropriate policies and procedures, including coordinated vulnerability disclosure policies to process and remediate potential vulnerabilities. The support period for which the manufacturer ensures the effective handling of vulnerabilities should be no less than five years, or the product's lifecycle, if shorter. When the manufacturer believes that the product does not comply with the CRA, it must take corrective action and even withdraw/recall that product from the market;
- Due diligence: Manufacturers will have to ensure that when integrating components sourced from third parties, those components do not compromise the cybersecurity of the product;

---

<sup>51</sup> For further analysis see also The CRA, explained - Cyber Resilience Act, Available at: <https://www.cyberresilienceact.eu/the-cra-explained/>

- Security updates: Manufacturers will have to maintain such updates available for a minimum of 10 years or for the remainder of the support period, according to the product's expected use, whichever is longer;
- Conformity assessment: Manufacturers will have to perform a third-party conformity assessment or conduct a self-assessment, for products that are not important or critical products under Annex III, to demonstrate conformity with the essential requirements. When a product with digital elements is certified under a European cybersecurity certificate at least at level 'substantial' no third-party conformity assessment is required;
- CE marking: The manufacturer will have to affix it to the product following the conformity assessment;
- EU declaration of conformity: Serving as a presumption of compliance, the manufacturer will have to keep it at the disposal of the national authorities for 10 years after the product has been placed on the market or for the support period, whichever is longer;
- User information and instructions: It refers to guidance on the safe installation and operation of the product. The manufacturer will have to keep it for 10 years or the support period, online or physically;
- Cooperation with market surveillance authorities: Manufacturers will have to keep the technical documentation and the EU declaration of conformity at the disposal of market surveillance authorities for at least 10 years or for the support period. They shall also provide the authorities with any other information and documentation requested;
- Reporting obligations: Manufacturers will have to report any actively exploited vulnerability or incidents impacting product safety to the computer security incident response team (CSIRT) designated as coordinator, and to ENISA within 24 hours. They will also have to inform market surveillance authorities and users affected by the reported incidents.

In turn, before placing imported products with digital elements on the EU market, the importers have to ensure that the manufacturers have fulfilled their obligations. Moreover, those importers will have to report cybersecurity risks and vulnerabilities to manufacturers, provide information to market surveillance authorities, keep product's documentation, and when needed, take further action for example through corrective measures or by withdrawing/recalling products which are not compliant with the essential requirements.

Distributors bear the responsibility of verifying that both manufacturers and importers have fulfilled their obligations to provide the technical information and instructions, the declaration of conformity and the CE marking. Similarly to distributors, they bear the duty of informing manufacturers and market surveillance authorities when they become aware of non-compliance to the essential requirements, take corrective measures or withdraw/recall the product from the EU market, and cooperate with market surveillance authorities for issuing all requested information and documentation.

### 6.3. Resulting requirements for FLUTE

The FLUTE platform is an open-source platform that utilizes armoured FL to enable the analysis of patient data across various hospitals and research centres. The CRA mainly focuses on non-embedded software, i.e., software that is additional to the primary function of the device on which it is downloaded<sup>52</sup> and

<sup>52</sup> European Commission, Study on Safety of non-embedded software; Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems. Available at: <https://op.europa.eu/en/publication-detail/-/publication/aad6a287-5523-11e9-a8ed-01aa75ed71a1/language-en>

exempts from its scope free and open-source software, as well as pure PaaS or SaaS, unless the latter is used to process remotely the data generated by a hardware product placed in the European market<sup>53</sup>. Moreover, the CRA requirements will become applicable of 11 December 2027, so after the end of the project.

Given this, the requirements of CRA would not apply to the platform during the lifetime of the project. Nonetheless, essential requirements of the CRA may serve as inspiration and benchmark e.g., ensure protection from unauthorized access, protect the confidentiality of stored data by encrypting them, commercialize the platform without exploitable vulnerabilities as elements incorporated in the development process of the platform, even though not formally bound by them. Such an approach can positively impact the market and regulatory acceptance of project's outcomes.

## 6.4. Cybersecurity Certification

### 6.4.1. Overview

With the entry into force of the EU Cybersecurity Act<sup>54</sup>, the European Union established rules for a harmonised cybersecurity certification framework as a comprehensive set of rules, technical requirements, standards and procedures that apply to the certification or conformity assessment of specific Information and communications technology (ICT) products, services or processes<sup>55</sup>. This voluntary framework (unless otherwise specified by EU or Member States law) results in a harmonised landscape across the EU, both in terms of security requirements for ICT solutions and the assessment methodology.

The European Union Agency for Cybersecurity (ENISA) holds a central role in the development of the schemes and in providing guidance to all interested. According to the EU Cybersecurity Act, ENISA shall facilitate the establishment and take-up of European and international standards for risk management and for the security of ICT products, services and processes. This supportive role includes aspects, such as monitoring standardisation developments and cybersecurity trends in the market, recommending appropriate technical specifications, preparing candidate European cybersecurity certification schemes taking into account existing schemes and standards, evaluating adopted European cybersecurity certification schemes, publishing guidelines and developing good practices, concerning the cybersecurity requirements, in cooperation with national cybersecurity certification authorities and industry, etc.

### 6.4.2. Purpose of EU cybersecurity certification

The cybersecurity certification under EU Cybersecurity Act serves a compliance tool for product manufacturers and service providers of the certified ICT products, services and processes. Moreover, it is a trust instrument. By achieving recognition of the certification at EU level vendors and service providers are able to reach more customers and demonstrate easier the cybersecurity of their solutions, in terms of resistance to certain levels of attacks. Thus, the focus is on achieving security objectives (e.g., protection of stored, transmitted or otherwise processed data against accidental or unauthorised destruction, loss or alteration)<sup>56</sup>. Having a trustworthy proof of cybersecurity robustness facilitates access to the market and gives a competitive advantage. Even in the case of solutions already certified with existing schemes, it is

<sup>53</sup> CRA guide for software developers - Cyber Resilience Act, <https://www.cyberresilienceact.eu/cra-guide-for-software-developers/>

<sup>54</sup> Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (Text with EEA relevance)

<sup>55</sup> According to the definition provided in Article 2 of the EU Cybersecurity Act.

<sup>56</sup> Article 51 EU Cybersecurity Act.

envisaged that ENISA and Member States will provide guidance to smoothen the transition process and compare requirements from existing schemes to the EU ones to facilitate transition. The ultimate goal of the cybersecurity certification is to empower the EU Digital Single Market<sup>57</sup>. Harmonisation of rules and procedures eliminates the multiplication of conflicting or overlapping national cybersecurity certification schemes and thus reduces operational costs for undertakings. Additionally, this uniformity prevents instances of ‘certification shopping’.

### 6.4.3. EU cybersecurity certification schemes

The certification is operationalized through the EU cybersecurity certification schemes. These schemes are prepared by ENISA upon request of the EC or Member States<sup>58</sup>. When preparing the scheme, ENISA should consult all relevant stakeholders (e.g., European standardisation organisations) by means of a formal, open, transparent and inclusive consultation process and shall establish an ad hoc working group of experts that supports in the development of a certain scheme. A draft scheme enters into force once it becomes part of the EU legislation through the adoption of an implementing act by the EC.

ENISA maintains a dedicated website providing information on current and past EU cybersecurity certification schemes, EU cybersecurity certificates and EU statements of conformity.<sup>59</sup>

Some important characteristics of certification schemes include:

- Coverage and scope: A scheme covers several aspects, such as: the subject matter and scope, specifying the type or categories of ICT products, services and processes covered; a clear description of the purpose of the scheme and of how the selected standards, evaluation methods and assurance levels correspond; references to the international, European or national standards applied in the evaluation, technical specifications or other cybersecurity specifications defined in the scheme; an indication of whether conformity self-assessment is permitted under the scheme; the specific evaluation criteria and methods used; the information that applicants shall make available to accredited conformity assessment bodies (CBs); rules for monitoring compliance of ICT products, services and processes with the requirements of the EU cybersecurity certificates or the EU statements of conformity, etc<sup>60</sup>.
- Risk-based approach: Certification schemes may specify one or more of the following assurance levels for ICT products, services and processes: ‘basic’, ‘substantial’ or ‘high’. The assurance level should correspond with the cybersecurity risk associated with the intended use of the ICT solution to be certified. Each assurance level is accompanied by different requirements specified in the EU Cybersecurity Act.
- Clearly indicated assessment process: The scheme should provide for clear and understandable means for consumers or other users to differentiate between those ICT products, services or processes subject to self-assessment, and those that are certified by a third party (for certifying entities see below).

The EC assesses the efficiency and use of the adopted EU cybersecurity certification schemes and takes a decision whether to make a specific scheme mandatory. The EC also publishes a Union rolling work

<sup>57</sup> ENISA, About EU Cyber Certification, Available at: [https://certification.enisa.europa.eu/about-eu-cyber-certification\\_en](https://certification.enisa.europa.eu/about-eu-cyber-certification_en), European Commission, The EU Cybersecurity Certification Framework, Available at: <https://digital-strategy.ec.europa.eu/en/policies/cybersecurity-certification-framework>

<sup>58</sup> For more information see also [https://certification.enisa.europa.eu/about-eu-certification/developing-certification-schemes\\_en](https://certification.enisa.europa.eu/about-eu-certification/developing-certification-schemes_en).

<sup>59</sup> EU Cybersecurity Certification, Available at: [https://certification.enisa.europa.eu/index\\_en](https://certification.enisa.europa.eu/index_en)

<sup>60</sup> Article 54 EU Cybersecurity Act.

programme for EU cybersecurity certification<sup>61</sup> that identifies strategic priorities for future European cybersecurity certification schemes. The EC may request ENISA to prepare a candidate scheme or to review an existing European cybersecurity certification scheme on the basis of the Union rolling work programme.

#### 6.4.4. EU cybersecurity certification process

Once the scheme is implemented, a manufacturer or provider of ICT products, services or processes may begin the certification process. Under an EU cybersecurity certification scheme it is possible to have either a third-party conformity assessment or self-assessment, or both.

- Third-party conformity assessment: CBs perform the certification<sup>62</sup> by auditing and/or testing and/or certifying ICT products, ICT services and ICT processes. CBs issue EU cybersecurity certificates referring to assurance level 'basic' or 'substantial'. However, national cybersecurity certification authorities, or a conformity assessment body upon prior approval by the national cybersecurity certification authority are responsible for issuing an EU cybersecurity certificate of assurance level 'high'.
- Conformity self-assessment: Manufacturers or providers of ICT solutions can carry out the conformity assessment themselves in the case of low complexity ICT products, services or processes that present a low risk to the public corresponding to assurance level 'basic', such as simple design and production mechanisms. The self-assessment should result in an EU statement of conformity stating that a specific ICT product, service or process complies with the requirements of the EU cybersecurity certification scheme at issue.

Certificates and EU statements of conformity have a duration specified in the relevant certification scheme; they can be renewed, provided that the requirements of the scheme continue to apply. They are recognised in all Member States and serve as a presumption of conformity of the certified ICT solutions with the scheme requirements.

#### 6.4.5. Existing or candidate schemes

EU cybersecurity certification scheme on Common Criteria (EUCC) is the first ever scheme which was adopted on 31 January 2024. It targets ICT products such as hardware and software products and components. ENISA published the state-of-the-art documents supporting the scheme as listed in its Annex 1. Two technical domains are currently widely used for certification at levels AVA\_VAN.4 and AVA\_VAN.5. The first technical domain is the 'Smart cards and similar devices', which may refer to Trusted Platform Modules as used in Trusted Computing<sup>63</sup>. This scheme does not allow for self-assessment. Instead, the Information Technology Security Evaluation Facilities (ITSEF) and CB should carry out the conformity assessment. CB shall issue EUCC certificates at assurance levels 'substantial' or 'high'. The validity of certificates should, in principle, not exceed 5 years.

<sup>61</sup> European Commission, Union Rolling Work Programme - EU Cybersecurity Certification, Available at: <https://digital-strategy.ec.europa.eu/en/library/union-rolling-work-programme-european-cybersecurity-certification>

<sup>62</sup> The conformity assessment bodies shall be accredited by national accreditation bodies appointed pursuant to Regulation 765/2008. Where a national cybersecurity certification authority issues EU cybersecurity certificates, the certification body of that national cybersecurity certification authority shall be accredited as a CB.

<sup>63</sup> Commission Implementing Regulation (EU) 2024/482 of 31 January 2024 laying down rules for the application of Regulation (EU) 2019/881 of the European Parliament and of the Council as regards the adoption of the European Common Criteria-based cybersecurity certification scheme (EUCC).

Candidate EU cybersecurity certification scheme on cloud services (EUCS) targets a specific category of ICT services and it is based on the ISO/IEC 17065 standard in terms of applicable requirements to conformity assessment bodies performing certification. The scheme is not final, as it will now enter the process of the European Cybersecurity Certification Group (ECCG) opinion. EUCS is intended to be a horizontal scheme, applying the same criteria to all cloud services, with three levels of assurance. These criteria apply to the design and implementation of the cloud service (e.g., security features, essential processes used throughout its lifecycle, in particular for development, deployment and operation). Users of the scheme may be cloud service providers; cloud service customers who wish to benefit from the evidence provided with certified cloud services to make informed decisions related to the security of these cloud services; regulatory authorities.

#### **6.4.6. Resulting requirements for FLUTE**

While not mandatory, cybersecurity certification may be a way for some of the project results to gain market recognition. For example, the technical domain ‘Smart cards and similar devices’ of the EUCC scheme, which may refer to Trusted Platform Modules as used in Trusted Computing, can be a potentially interesting alley for certification. However, cybersecurity certification schemes are quite recent as a framework. The fact that for the moment there is no available scheme on AI/FL makes the possibility of certification harder. Consortium partners should closely follow the relevant developments to observe when circumstances can be more favourable for the certification of project outcomes.

## 7. European Health Data Space

### 7.1. Overview

The European Health Data Space Regulation<sup>64</sup> (EHDSR) was officially adopted and published on 5 March 2025. It is the most relevant EU act concerning health data.

The overarching purpose of the EHDSR is to strengthen patients' rights to their health data and to open up databases containing electronic health data (see definition below), to make better use of it, both for the patients and the larger community.

Specific goals set by the EHDSR include:

- reinforcing the rights of natural persons (patients) in relation to the availability and control of their electronic health data;
- providing common rules and mechanisms for primary use of electronic health data and secondary use of electronic health data;
- laying down harmonized requirements for two mandatory harmonised software components of electronic health records (EHR) systems on the EU market;
- establishing mandatory cross-border infrastructure enabling the primary and secondary use of electronic health data across the EU.

EHDS builds upon two pillars:

- **Primary use of electronic health data:** Use of data in the context of healthcare. This will cover use of data for treating the patient, but also the prescription and dispensation of medicinal products and medical devices, as well as for relevant social security, administrative or reimbursement services.
- **Secondary use of electronic health data:** Use of data for other purposes that benefit the society. Examples: research, innovation, policy-making, patient safety, personalised medicine, official statistics or regulatory activities. The second pillar i.e. rules on secondary use of electronic health data, is of relevance to research projects funded by the EU, such as FLUTE, both from the perspective of obtaining data for research (as health data users) and from the perspective of making the repository data available for further research (as health data holder). Thus, below we focus on main principles of the EHDS, to the extent that it relates to FLUTE project.

### 7.2. Key terms in EHDSR

Below we explain the key terms for understanding the EHDSR.

- **Electronic health data (EHD):** Term covers personal or non-personal electronic health data. Potentially all health-related data (both anonymized or pseudonymized) included in the FLUTE repository would be considered as electronic health data.
- **Health data access body (HDAB):** Authorities set up by EU countries which will play a central role in making EHD available for secondary purposes.

---

<sup>64</sup> Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847

- Health data holder: Person or entity who has the right or obligation in accordance with the law, and in its capacity as a (joint) controller to process personal EHD for the purposes of provision of healthcare, research (and others) OR ability to make non-personal EHD available through the control of the technical design of a product and related services. Examples: hospital, research organization, developer of a wellness app.
- Health data user: Person or entity who has been granted lawful access to personal or non-personal EHD for secondary use pursuant to a data permit or approval, includes EU institutions.

### 7.3. Access to data for secondary use according to the EHDS

EHDS establishes a clear framework for accessing EHD for secondary use, including among other purposes:

- education or teaching activities in health or care sectors at vocational or higher education level;
- scientific research related to health or care sectors that contributes to public health or health technology assessments, or ensures high levels of quality and safety of healthcare, of medicinal products or of medical devices, with the aim of benefiting end-users, such as patients, health professionals and health administrators, including: (i) development and innovation activities for products or services; (ii) training, testing and evaluation of algorithms, including in medical devices, in vitro diagnostic medical devices, AI systems and digital health applications;
- improvement of the delivery of care, of the optimisation of treatment and of the provision of healthcare, based on the electronic health data of other natural persons.

However, the regulation also lists certain uses which are prohibited, for example:

- taking decisions that produce legal, social or economic effects detrimental to a natural person or a group based on their EHD;
- taking decisions in relation to a natural person in relation to job offers, offering less favourable terms in the provision of goods or services, or taking any other decisions which result in discrimination on the basis of the EHD;
- carrying out advertising or marketing activities;
- developing products or services that may harm individuals, public health or society at large, such as illicit drugs, alcoholic beverages, tobacco and nicotine products, weaponry or products or services which are designed or modified in such a way that they create addiction, contravene public order or cause a risk for human health;
- carrying out activities in conflict with ethical provisions laid down in national law;

Standardized rules will enable researchers (as potential data users) to request various types of health-related data, in a standardized manner, as depicted in the schema shown below<sup>65</sup>.

<sup>65</sup> Adapted from presentation EUHPP webinar on European Health Data Space (2/3): Secondary Use of health data (27 February 2025) of EC. Available at: [https://health.ec.europa.eu/latest-updates/recording-and-presentation-euhpp-webinar-european-health-data-space-23-secondary-use-health-data-27-2025-03-03\\_en](https://health.ec.europa.eu/latest-updates/recording-and-presentation-euhpp-webinar-european-health-data-space-23-secondary-use-health-data-27-2025-03-03_en)

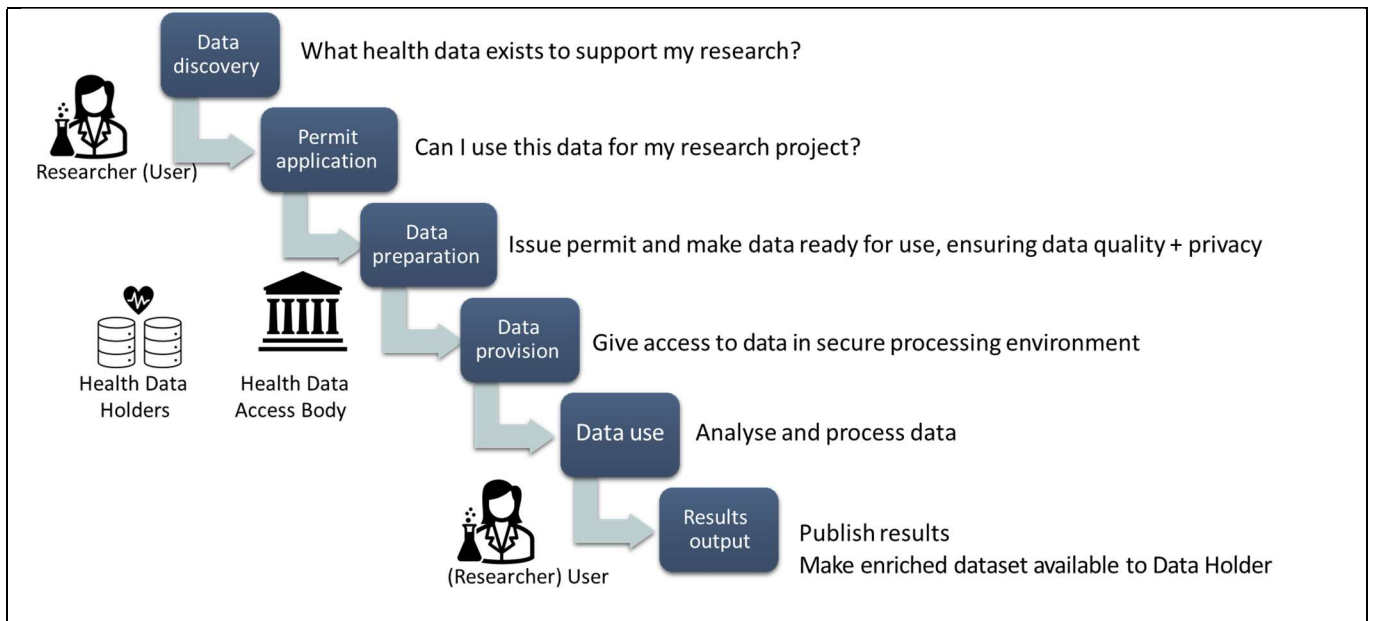


Figure 1 Summary of access to EHD for secondary purposes

In more detail, health data holders will be obligated to inform HDAB about their datasets and their characteristics. Furthermore, they will need to make their EHD available for re-use on request of HDAB. Holders will be able to request reasonable fees for making their EHD available for secondary use. However, such fees must be objectively justified, transparent and proportionate to the costs of collecting and making the data available for secondary use.

Health data access bodies (HDAB) will be authorities set up by Member States. HDAB will keep a national dataset catalogue with a list of available datasets. National catalogues will be connected by EU datasets catalogue. Data catalogues will make the health data discoverable to potential data users and HDAB will inform the public about conditions under which electronic health data is made available. These bodies will examine the applications from potential users and issue data permits (i.e. administrative decisions which allow a health data user to access EHD) and approve data requests (requests for aggregated information). HDAB also will collect the requested EHD from the relevant health data holder and pre-process it (e.g. anonymize it or combine it with other datasets) and make it available for access of the health data user through a secure processing environment. If during the research there is a finding that may impact on the health of a natural person, the HDAB may inform the natural person and his or her treating health professional about that finding. Finally, HDAB will enforce the obligations of the EHDSR, for instance including power to fine a data holder which does not provide their datasets for secondary use.

Researchers who wish to access EHD and become health data users will have to request access at the HDAB. To do so, they will need to apply for the issuance of a data permit, but submitting an application and providing the required details of the proposed research.

HDAB examine the application and issue the permit or refuse it. Data permit will set out general conditions that apply to the user. Data users will have the right to access and process the electronic health data in accordance with the data permit delivered to them on the basis of EHDS. Data users will have to also make sure to publish the results of their research acknowledging the electronic health data sources and the fact that electronic health data has been obtained in the context of the EHDS. If the data is enriched, the dataset with such improvements and a description of the changes will be made available free of charge

to the original data holder. They will also have to inform the HDAB of any clinically significant findings that may influence the health status of the natural persons whose data is included in the dataset.

#### 7.4. Data quality and utility labels

As mentioned above, under the EHDSR the health data holders will need to inform health data access bodies about their datasets and their characteristics. Holders may also provide a Union data quality and utility label on their datasets, if they fulfil principles defined by the regulation and delegated acts, which will be adopted in the next years. For some data sets (e.g. those created with public funding, such as EU Horizon projects), adherence to those principles will be mandatory. In such case, the holder should have 'sufficient documentation' for the health data access body to confirm the accuracy of the label.

In particular, the data quality and utility label shall comply with the following elements:

- for data documentation: metadata, support documentation, the data dictionary, the format and standards used, the source of the data and, where applicable, the data model;
- for assessment of technical quality: the completeness, uniqueness, accuracy, validity, timeliness and consistency of the data;
- for data quality management processes: the level of maturity of the data quality management processes, including review and audit processes, and bias examination;
- for assessment of coverage: the period, population coverage and, where applicable, representativity of the population sampled, and the average timeframe in which a natural person appears in a dataset;
- for information on access and provision: the time between the collection of the electronic health data and their addition to the dataset and the time needed to provide electronic health data following the issuing of a data permit or a health data request approval;
- for information on data modifications: merging and adding data to an existing dataset, including links with other datasets.

The EC is empowered to adopt delegated acts to amend this list. Also, the EC can issue implementing acts to set out 'visual characteristics and technical specifications of the data quality and utility label', based on the elements referred above.

#### 7.5. Secure processing environments

Data for secondary use can only be provided in anonymized format or in pseudonymized format (only if the purpose of the data user's processing cannot be achieved with anonymised data). The information necessary to reverse the pseudonymisation will be available only to the HDAB. Moreover, secondary use access to the requested electronic health data will be done through a secure processing environment, with technical and organisational measures and security and interoperability requirements.

Secure processing environment<sup>66</sup> is a physical or virtual environment and organisational means to ensure compliance with Union law, such as GDPR, in particular with regard to data subjects' rights, intellectual property rights, and commercial and statistical confidentiality, integrity and accessibility, as well as with applicable national law, and to allow the entity providing the secure processing environment to determine and supervise all data processing actions, including the display, storage, download and export of data and the calculation of derivative data through computational algorithms.

<sup>66</sup> Definition from Data Governance Act, Article 2 (20)

According to the EHDSR, secure processing environment should<sup>67</sup>:

- restrict access to the secure processing environment to authorised persons listed in the respective data permit;
- minimise the risk of the unauthorised reading, copying, modification or removal of electronic health data hosted in the secure processing environment through state-of-the-art technological means;
- limit the input of EHD and the inspection, modification or deletion of EHD in the secure processing environment to a limited number of authorised identifiable individuals;
- ensure that data users have access only to the electronic health data covered by their data permit, by means of individual and unique user identities and confidential access modes only;
- keep identifiable logs of access to the secure processing environment for the period of time necessary to verify and audit all processing operations in that environment;
- ensure compliance and monitor the security measures referred to in Article 50 of EHDS to mitigate potential security threats.

The data users will only be able to download *non-personal* electronic health data from the secure processing environment. The EC will, by means of implementing acts, provide for the technical, information security and interoperability requirements for the secure processing environments.

## 7.6. Resulting requirements for FLUTE

The EHDSR was published in the Official Journal on 5 March 2025 and entered into force 20 days later. However, it will start to apply in a phased way, Chapter IV on secondary use applying to most categories of EHD by 26 March 2029. For some categories, such as genetic data, it will apply from 6 years after entry into force, so by 26 March 2031. Given the lifetime of the project, EHDS secondary use provisions will not directly apply to results of FLUTE.

Nevertheless, the future operationalization of EHDS should be considered when discussing exploitation and sustainability plans for the FLUTE platform and collected data, in particular:

- FLUTE platform may potentially be developed in the future to become a ‘secure processing environment’ or otherwise a tool to assist health data holders or HDAB in providing access to data in a EHDS compliant manner.
- Data Providers should consider quality label requirements when collecting data.
- In the context of potential FLUTE data access and governance process, the framework of EHDS should be taken into account.
- Identified EHDS patient documents linked to the FLUTE project are EPS (European Patient Summary), Hospital Discharge Letter and European Lab Report
- FLUTE standardization partner, HL7, has also started standardization work of cancer data in the context of the EHDS

---

<sup>67</sup> See Article 50 EHDS Regulation.

## 8. Data Governance Act

### 8.1. General information

The goal of the Data Governance Act<sup>68</sup> (DGA), which is a part of the European Data Strategy, is to provide a framework to enhance trust in voluntary data sharing for the benefit of businesses and citizens. The regulation entered into force on 24 September 2023 and is directly applicable in all EU Member States.

Key areas of regulation<sup>69</sup>:

- Re-use of certain categories of data held by public sector bodies: facilitating re-use of protected data (e.g. personal data and commercially confidential data) held by the public sector.
- Data intermediation services: set of rules for providers of data intermediation services (such as data marketplaces) to ensure that they will function as trustworthy organisers of data sharing or pooling within the common European data spaces.
- Data altruism: is about individuals and companies giving their consent or permission to make available data that they generate – voluntarily and without reward – to be used in the public interest.
- European Data Innovation Board: to facilitate the sharing of best practices, in particular on data intermediation, data altruism and the use of public data that cannot be made available as open data, as well as on the prioritisation of cross-sectoral interoperability standards.

Both personal and non-personal data are in scope of the DGA, and wherever personal data is concerned, the GDPR applies. DGA refers to European data spaces<sup>70</sup> and future specific rules applicable to them. EHDS (described above) is to be an example of such data space. Last but not least, DGA introduces the concept of secure processing environment<sup>71</sup> (see also Section 7.5).

Below, we describe the guidelines related to data intermediation services and data altruism, as most relevant for FLUTE and its future sustainability planning.

### 8.2. Data Intermediation Services – definition and types

It has been noted that there are various economic functions which data intermediaries could assume and hence they may be promising to policymakers:

- Perform a matching function that can bring together data holders and users;
- Improve the accessibility of data;
- Decrease information asymmetries regarding data and reduce costs of risk for actors in the data ecosystem to share their data with others;
- Can reduce transaction costs, e.g. by standardisation and technical and contractual management of data transfers and enforcement of the agreed conditions;

<sup>68</sup> Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance)

<sup>69</sup> Based on European Commission, Data Governance Act explained, Available at: <https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained>

<sup>70</sup> See Recital 2 DGA.

<sup>71</sup> Under Article 1(2) secure processing environment means the physical or virtual environment and organisational means to provide the opportunity to re-use data in a manner that allows for the operator of the secure processing environment to determine and supervise all data processing actions, including to display, storage, download, export of the data and calculation of derivative data through computational algorithms.

- can capture the value of network effects and pass them on to data holders and users<sup>72</sup>.

However, the general term ‘data intermediary’ should not be confused with the legally recognised concept of ‘data intermediation services’ (DIS) as laid down in the DGA. Under the DGA, ‘data sharing’ means the provision of data by a data subject or a data holder to a data user for the purpose of the joint or individual use of such data, based on voluntary agreements or Union or national law, directly or through an intermediary, for example under open or commercial licences subject to a fee or free of charge. Furthermore, ‘data intermediation service’ means a service which aims to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other, through technical, legal or other means, including for the purpose of exercising the rights of data subjects in relation to personal data, excluding at least the following:

- Services that obtain data from data holders and aggregate, enrich or transform the data for the purpose of adding substantial value to it and license the use of the resulting data to data users, without establishing a commercial relationship between data holders and data users;
- Services that focus on the intermediation of copyright-protected content;
- Services that are exclusively used by one data holder in order to enable the use of the data held by that data holder, or that are used by multiple legal persons in a closed group, including supplier or customer relationships or collaborations established by contract, in particular those that have as a main objective to ensure the functionalities of objects and devices connected to the Internet of Things;
- Data sharing services offered by public sector bodies that do not aim to establish commercial relationships.

In essence, the main characteristics of a DIS are: (i) establishing commercial relationships for data sharing, (ii) between an undetermined number of data subjects/data holders and data users, (iii) through technical, legal or other means. There are various categories of DIS, as explained below.

Table 5 Categories of DIS under DGA

<b>Art. 10 (a)</b> <b>DIS between data holders – potential data users</b>	<b>Art. 10 (b)</b> <b>DIS between data subjects or natural persons – potential data users</b>	<b>Art. 10 (c)</b> <b>Services of data cooperatives</b>
--	--	--

<sup>72</sup> Based on Heiko Richter, Looking at the Data Governance Act and Beyond: How to Better Integrate Data Intermediaries in the Market Order for Data Sharing, GRUR International, Volume 72, Issue 5, May 2023, Pages 458–470, <https://doi.org/10.1093/grurint/ikad014> and further sources cited there.

<p>Data exchange</p> <ul style="list-style-type: none"> <li>• Bilateral or multilateral exchanges of data</li> <li>• Creation of platforms or databases enabling the exchange or joint use of data</li> </ul> <p>Establishment of other specific infrastructure for the interconnection of data holders with data users</p>	<p>Data availability</p> <ul style="list-style-type: none"> <li>• Data subjects (individuals): seek to make their personal data available or</li> <li>• Legal persons: seek to make non-personal data available</li> </ul> <p>Enabling the exercise of the data subjects' rights (GDPR)</p>	<p>Organisations which support their members (data subjects, one-person undertakings or SMEs) in exercising their rights in respect to data:</p> <ul style="list-style-type: none"> <li>• Informed consent choices</li> <li>• Exchanging views on data processing purposes</li> </ul> <p>Negotiating favourable terms and conditions on behalf of its members</p>
<p><b>Examples:</b> Data marketplaces, orchestrators of data sharing ecosystems, data pools</p>	<p><b>Examples:</b> Personal Information Management Systems (PIMS) <a href="https://digi.me/">https://digi.me/</a>, <a href="https://citizenme.com/">https://citizenme.com/</a></p>	<p><b>Examples:</b> Driver's Seat, Swash</p>

### 8.3. Requirements for providing DIS

DGA specifies the requirements for the DIS with the aim to achieve the trustworthy provision of DIS. Neutrality and independence are regarded as key elements to bring about more trust and control. Hence, DIS providers (DISPs) must observe a number of requirements when providing the service. They include<sup>73</sup>:

- DISP's independence (structural separation):
  - DISP must provide data intermediation services through a separate legal person;
  - DISP's commercial terms, including pricing, cannot be dependent upon use of other services provided by the provider or by a related entity;
  - DISP may offer additional specific tools and services to data holders or data subjects for the specific purpose of facilitating the exchange of data, such as temporary storage, curation, conversion, anonymisation and pseudonymisation. Those tools can only be used at the explicit request of the data holder or data subject;
- Restrictions on the use of data:
  - DISP cannot use the data for which it provides data intermediation services for purposes other than to put them at the disposal of data users;
  - DISP is limited in using the data collected for the purpose of provision of the services (holder/user data); they may only use such data for the development of the data intermediation service, including detection of fraud or cybersecurity;
  - DISP must facilitate the exchange of the data in the format in which it receives it from a data subject or a data holder;

<sup>73</sup> See Article 12 DGA.

- DISP can convert the data into specific formats only to enhance interoperability within and across sectors or if requested by the data user or required by law; opt out possibility must be available;
- Service conditions: transparency and fairness
  - DISP must ensure that the procedure for access to its service is fair, transparent and non-discriminatory for both data subjects and data holders, as well as for data users, including with regard to prices and terms of service (neutrality of the DISP)
  - Prohibition against DISPs making their terms 'dependent upon whether or to what degree the data holder or data user uses other services provided by the same provider or a related entity' (prohibition of bundling services)
- Legal, technical and organizational safeguards:
  - DISP must maintain a log record of the data intermediation activity;
  - DISP must put in place adequate technical, legal and organisational measures in order to prevent the transfer of or access to non-personal data that is unlawful;
  - DISP must inform data holders in the event of an unauthorised transfer, access or use of the non-personal data that they shared;
  - DISP must take necessary measures to ensure an appropriate level of security for the storage, processing and transmission of non-personal data and highest level of security when storing or sharing competitively sensitive information;
  - DISP must have procedures in place to prevent fraudulent or abusive practices in relation to parties which want to access the data;
  - in the event of its insolvency, DISP must ensure a reasonable continuity of the provision of its data intermediation services, including for the storage of data.
- Interoperability:
  - DISP must take appropriate measures to ensure interoperability with other data intermediation services, e.g. by means of commonly used open standards;
  - the EC encourages and facilitates Union-wide codes of conduct, especially on interoperability, and '[t]he European Data Innovation Board should facilitate the emergence of additional industry standards, where necessary.

DISPs which offer services to data subjects must act in the data subjects' best interest where they facilitate the exercise of their rights [Article 12(m)]. Additional requirements apply to DISPs which provide tools for obtaining consent from data subjects or permissions to process data made available by data holders.

DGA installs a mandatory compliance regime which requires DISPs to officially register their services as a precondition for lawfully providing them in the EU. Under DGA, a provider which intends to provide the DIS must submit a notification to the competent authority in the Member State where it is established. DISPs located outside the EU will have to appoint a representative in the EU<sup>74</sup>.

---

<sup>74</sup> See Article 11.1-4 DGA.

A DISP which obtained confirmation from the authority that it provides the services in accordance with the DGA requirements may use the label of ‘data intermediation services provider recognised in the Union’ as well as a common logo.



EU Recognised  
Data Intermediary

Figure 2 Data intermediation services provider recognised in the Union label



EU Recognised  
Data Intermediary

DISPs are subject to monitoring of compliance by competent authorities established in each Member State<sup>75</sup>. If the DISPs do not comply, they face penalties and can be suspended from offering their services. DGA<sup>76</sup> requires that penalties be effective, proportionate and dissuasive and leaves it to the Member States to specify the penalties by outlining some criteria to be taken into account when imposing penalties. Data holders and data users can assert contractual claims against DISPs.

#### 8.4. Data altruism under the DGA

Data altruism means voluntary sharing of data

- on the basis of the consent of data subjects to process personal data pertaining to them, or permissions of data holders to allow the use of their non-personal data,
- without seeking or receiving a reward that goes beyond costs they incur where they make their data available and
- for objectives of general interest as provided for in national law<sup>77</sup>.

Those objectives may include, where applicable, for example healthcare, combating climate change, improving mobility, facilitating the development, production and dissemination of official statistics, improving the provision of public services, public policy making or scientific research purposes in the general interest.

While DGA does not preclude national policies for data altruism, it sets down rules for recognised data altruism organisations. Organisations may apply for registration in the public national register of recognised data altruism organisations and use the label ‘data altruism organisation recognised in the Union’ in its written and spoken communication, as well as a common logo.

<sup>75</sup> See Article 14 DGA.

<sup>76</sup> See Article 34 DGA.

<sup>77</sup> Modified from definition in Article 1(16) DGA.



Figure 3 Data altruism organisation recognised in the Union labels

An organisation which carries out data altruism activities must fulfil a number of requirements to qualify for registration in a public national register of Recognized Data Altruism Organizations (RDAOs), in particular, it must:

- operate on a not-for-profit basis;
- be a legal person established according to the national law and be legally independent from any entity that operates on a for-profit basis;
- carry out its data altruism activities through a structure that is functionally separate from its other activities;
- comply with the rulebook adopted by the EC<sup>78</sup>.

Furthermore, during its operation a RDAO must also observe certain requirements, such as:

- comply with transparency requirements, e.g., keep full and accurate records concerning the persons or entities providing data, the duration of the processing of this data, the purpose of it and the fees paid<sup>79</sup>;
- file an annual activity report with the competent authority<sup>80</sup>;
- inform the data subject and data holders about, *inter alia*, objectives of general interest, purpose for which their data will be processed<sup>81</sup> and tools for obtaining consent or permission for processing and withdrawing it<sup>82</sup>;
- take measures to ensure an appropriate level of security for the storage and processing of non-personal data that it has collected based on data altruism<sup>83</sup> and inform data holders about unauthorized disclose or breach of their non-personal data<sup>84</sup>;
- refrain from using the entrusted data for other objectives than those of general interest for which the data subject or data holder allows the processing;
- refrain from using misleading marketing practices to solicit the provision of data.<sup>85</sup>

<sup>78</sup> See Article 18 DGA.

<sup>79</sup> See Article 20.1 DGA.

<sup>80</sup> See Article 20.2 DGA.

<sup>81</sup> See Article 21.1 DGA.

<sup>82</sup> See Article 21.3 DGA.

<sup>83</sup> See Article 21.4 DGA.

<sup>84</sup> See Article 21.5 DGA.

<sup>85</sup> See Article 21.2 DGA.

In order to facilitate the collection of data based on data altruism, EC will be empowered to adopt a European data altruism consent form.<sup>86</sup>

The table below illustrates the main differences between DISP and RDAO.

*Table 6 The main differences between DISP and RDAO*

DISPs	RDAOs
<ul style="list-style-type: none"> <li>• Establish commercial relationships for the purposes of data sharing between data holders and data users,</li> <li>• DISP must notify competent authority &amp; may use the label/logo,</li> <li>• Strict requirements apply:               <ul style="list-style-type: none"> <li>○ Neutrality (e.g. separate entity)</li> <li>○ EU rep for non-EU DISP</li> <li>○ Data security etc.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Voluntary sharing of data by individuals or data holders for objectives of general interest w/o remuneration,</li> <li>• RDAO may register with authority &amp; may use the label/logo,</li> <li>• Legal requirements for registered organisations.</li> <li>• Use of European data altruism consent form &amp; comply with rulebook.</li> </ul>

### 8.5. Resulting requirements for FLUTE

The requirements of DGA need to be observed during implementation of the project:

- Consider DGA requirements when planning sustainability of the FLUTE platform,
- If the goal of the project is to create a DIS ('data marketplace'), the Platform must meet the DGA technical (e.g. logs, security) and interoperability requirements.
- It can be suspected that the rules of the 'secure processing environments' will likely become a benchmark for environments in which data is made available for research. For instance, under Article 40, where data altruism organisations process personal electronic health data using a secure processing environment, such environments shall also comply with the requirements set out in Article 50 of this Regulation. As such, they should be carefully considered in the context of the FLUTE platform design and features.
- On the interoperability aspect, HL7 has started some work on shareable and interoperable research artefacts.

<sup>86</sup> See Article 25 DGA.

## 9. Medical device regulation

### 9.1. Software as a Medical Device

Medical Device Regulation<sup>87</sup> (MDR) is applicable to software that is classified as a medical device. Therefore, understanding when a software amounts to a medical device and falls within the scope of the MDR is of crucial importance. For this purpose, we examine below the concept of software as a medical device under the MDR.

Under MDR, ‘medical device’ means any instrument, apparatus, appliance, **software**, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:

- diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,
- investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,
- providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations,

and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means<sup>88</sup>.

According to recital 19 of MDR it is necessary to clarify that software in its own right, when specifically intended by the manufacturer to be used for one or more of the medical purposes set out in the definition of a medical device, qualifies as a medical device, while software for general purposes, even when used in a healthcare setting, or software intended for life-style and well-being purposes is not a medical device. The qualification of software, either as a device or an accessory, is independent of the software’s location or the type of interconnection between the software and a device.

Hence, medical device software (MDSW) is software that is intended to be used, alone or in combination, for a purpose as specified in the definition of a ‘medical device’ in the MDR<sup>89</sup>. The term ‘intended purpose’ used in the definition is of significance, as the intention of the manufacturer plays a vital role in determining and distinguishing a medical device from other devices. The device’s ‘intended purpose’ refers to the manufacturer’s stated use of the device, which could be indicated amongst others means, on the label, instructions for use, or promotional or sales materials of the device. In Case C-219/11 *Brain Products*<sup>90</sup> the Court found that for software to fall within the scope of Directive 93/42 [legislation preceding MDR] ‘it is not sufficient that it [the software] be used in a medical context, but that it is also necessary that the intended purpose, defined by the manufacturer, is specifically medical.’ Further clarification on the qualification of software as a medical device is provided by the Medical Device Coordination Group (MDCG). MDCG has published guidelines<sup>91</sup> (MDCG Guidelines) that assist manufactures to ask relevant

<sup>87</sup> Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC

<sup>88</sup> Article 2(1) MDR.

<sup>89</sup> MDSW can also be software specified in In Vitro Diagnostic Medical Devices Regulation (IVDR).

<sup>90</sup> Case C-219/11 *Brain Products* Judgment of the Court (Third Chamber) of 22 November 2012, para 17

<sup>91</sup> MDCG 2019-11 Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR (October 2019).

questions in order to determine whether their software qualifies as MDSW under the MDR. According to the MDCG Guidelines software should be understood as a set of instructions that processes input data and creates output data.

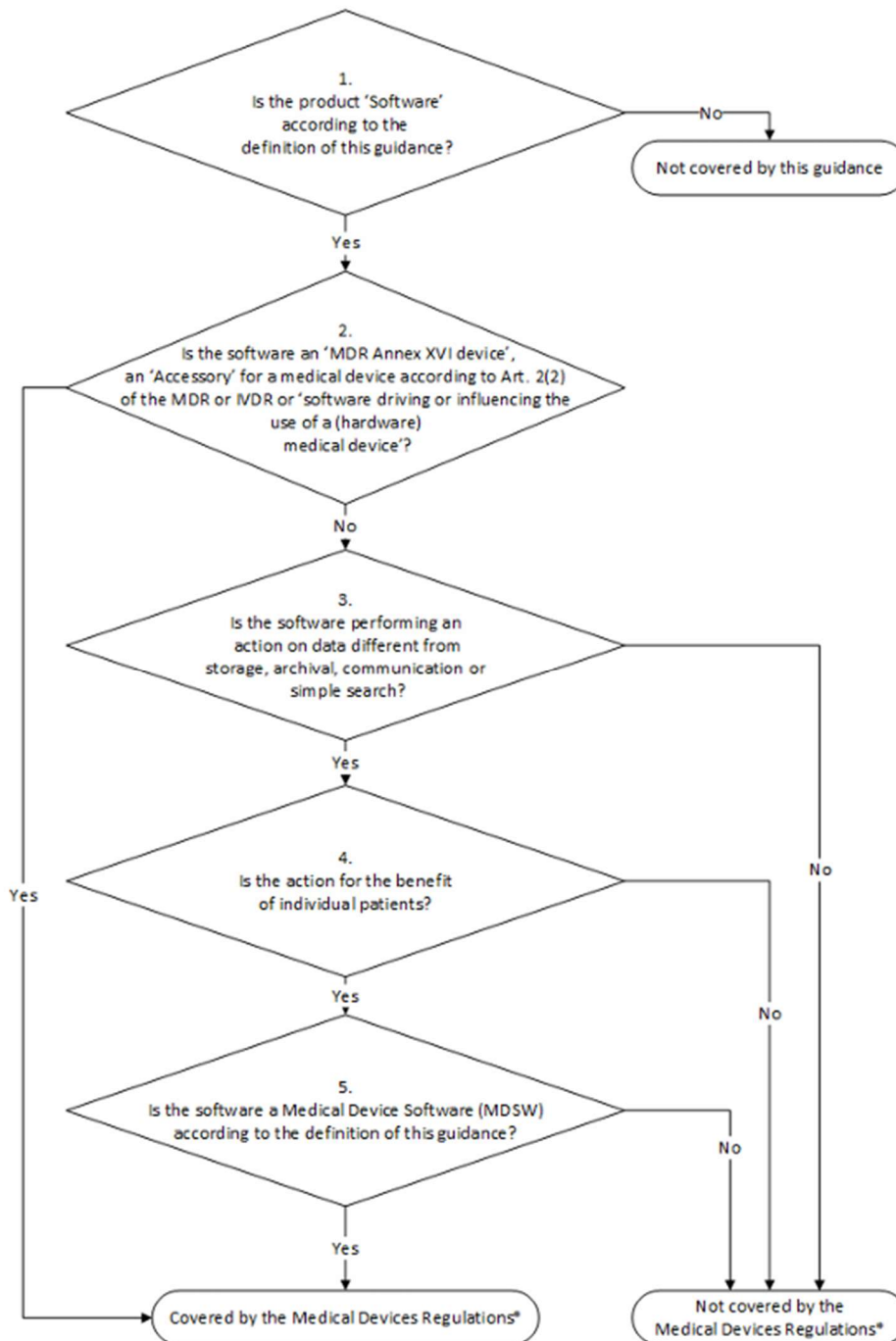


Figure 4 Qualification of MDSW<sup>92</sup>

<sup>92</sup> Source of figure 5: MDCG Guidelines.

Medical devices are subject to classification rules under MDR. Medical devices are divided into classes I, IIa, IIb and III, taking into account their intended purpose and their inherent risks. As a general rule, the greater the potential impact on patient health and the associated risk, the higher the classification class. Specific rules of classification of the medical devices are specified in Annex VIII of MDR.

When it comes to MDSW, the general rule in the Annex VIII states that software, which drives a device or influences the use of a device, shall fall within the same class as the device but if the software is independent of any other device, it shall be classified in its own right. Furthermore, Rule 11 states that software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause (1) death or an irreversible deterioration of a person's state of health, in which case it is in class III; or (2) a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb. Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb. All other software is classified as class I.

## 9.2. Resulting requirements for FLUTE

The applicability of MDR varies depending on the specific result of the project. In particular:

- The aim of the FLUTE project is to develop a dual-sided FLUTE platform which will be integrated with health data hubs located in four different countries. The FLUTE platform will make the siloed datasets available to researchers under privacy protection that complies with GDPR requirements, in standard FL settings and AI model development workflows. The key end users of the platform are researchers and innovators (on the demand side) and data hubs (on the supply side). The primary goal of the platform is to enable the analysis of patient data across various hospitals and research centres to develop an AI model for the prediction and diagnosis of prostate cancer. It is at a later stage that potentially an AI model trained with the use of the FLUTE platform will be used by healthcare practitioners as a prediction tool for individual patients. Therefore, the FLUTE platform does not serve directly a medical purpose in the meaning of the MDR and based on the above description and the MDCG guidelines, the FLUTE platform cannot be qualified as MDSW.
- Another result of the project is a novel federated AI toolset for diagnosis of csPCa. The impact of AI in improving information from medical images and creation of predictive models, will represent a significant advance in optimizing diagnosis of PCa and predicting its aggressiveness. If it is intended that the developed AI toolset would be that of facilitating the diagnosis of CsPCa in individual cases, such software can be classified as a medical device under, at least, class IIa, according to Rule 11, if it were intended to provide information which is used to take decisions for diagnosis or therapeutic purposes.
- For medical devices classified as class IIa and higher, the MDR specifies an appropriate regulatory process, including but not limited to the following obligations for manufacturers: establishing, documenting, implementing, and maintaining a risk management system; implementing a quality management system; conducting a clinical evaluation; continuously updating and improving the quality management system; ensuring the existence of technical documentation; undergoing a conformity assessment procedure by an EU Notified Body; obtaining an EU Declaration of Conformity and appropriate certificates; registering the device; and conducting post-market surveillance. Classification of a medical devices as class IIa or higher would also trigger classification of an AI model used therein

as high risk as described in Sections 5.4. These requirements will need to be taken into account in the exploitation plans of the partners which develop the AI toolset.

## 10. Conclusions

The initial iteration of D6.1, shared with project partners at M12, aimed to outline the fundamental legal and ethical considerations relevant to the FLUTE project. Partners were encouraged to seek guidance from the legal and ethics partner, TLX, whenever legal or ethical questions emerged during the project's execution. Additionally, TLX took a proactive role by organizing legal workshops and keeping partners informed about evolving issues.

Over the next 12 months, the document was updated to align with the project's development and changes in the legal landscape, particularly regarding the AIA, cybersecurity framework, and medical device regulation. Consequently, this final version of D6.1 provides project partners with comprehensive guidance on privacy and data protection, legal obligations for trustworthy AI, MDR compliance, and data governance. The document further elaborates on the application of key data protection principles within the FLUTE platform.

Looking ahead, WP6 will continue working closely with partners to maintain alignment with evolving legal and ethical frameworks in the context of health data sharing for the purpose of secure federated learning, including the EHDS, GDPR and DGA. We will also contribute to D6.2 GDPR guidelines and regulatory liaising. In particular, we will analyze the organizational measures that operators of federated learning platforms and data hubs should put in place to achieve and demonstrate regulatory compliance under GDPR, DGA, EHDS.

## 11. References

### Legal acts

- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC
- Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC
- Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)
- Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (Text with EEA relevance)
- Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance)
- Commission Implementing Regulation (EU) 2024/482 of 31 January 2024 laying down rules for the application of Regulation (EU) 2019/881 of the European Parliament and of the Council as regards the adoption of the European Common Criteria-based cybersecurity certification scheme (EUCC).
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024
- Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act) (Text with EEA relevance)
- Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847 (Text with EEA relevance)

### Guidelines

- AEPD (Spanish DPA), 'Synthetic data and data protection', (November 2023). Available at: <https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-data-and-data-protection>
- Article 29 Working Party, Opinion 05/2014 of the Article 29 Working Party on Anonymisation Techniques. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)
- Article 29 Working Party, Opinion 6/2013 on open data and public sector information ('PSI') reuse. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp207_en.pdf).

- Article 29 Working Party, Opinion 5/2014 on Anonymization Techniques. Available at: [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf).
- CNIL, Self-assessment guide for artificial intelligence (AI) systems, Available at: <https://www.cnil.fr/en/self-assessment-guide-artificial-intelligence-ai-systems>
- EDPB Guidelines 01/2025 on Pseudonymisation, Available at: [https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation\\_en](https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation_en)
- EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR, Available at: [https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr\\_en](https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor-gdpr_en)
- EDPB, Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research, 2 February 2021. Available at: [https://edpb.europa.eu/sites/default/files/files/file1/edpb\\_replyec\\_questionnaireresearch\\_final.pdf](https://edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf)
- ENISA, About EU Cyber Certification, Available at: [https://certification.enisa.europa.eu/about-eu-cyber-certification\\_en](https://certification.enisa.europa.eu/about-eu-cyber-certification_en)
- ENISA, Developing Certification Schemes, Available at: [https://certification.enisa.europa.eu/about-eu-certification/developing-certification-schemes\\_en](https://certification.enisa.europa.eu/about-eu-certification/developing-certification-schemes_en)
- European Commission, The EU Cybersecurity Certification Framework, Available at: <https://digital-strategy.ec.europa.eu/en/policies/cybersecurity-certification-framework>
- European Commission, JRC Publications Repository - Harmonised Standards for the European AI Act, Available at: [https://publications.jrc.ec.europa.eu/repository/bitstream/JRC139430/JRC139430\\_01.pdf](https://publications.jrc.ec.europa.eu/repository/bitstream/JRC139430/JRC139430_01.pdf)
- European Commission, Study on Safety of non-embedded software; Service, data access, and legal issues of advanced robots, autonomous, connected, and AI-based vehicles and systems. Available at: <https://op.europa.eu/en/publication-detail/-/publication/aad6a287-5523-11e9-a8ed-01aa75ed71a1/language-en>
- European Commission, Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act). Available at: <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act>.
- European Commission, Guidelines on AI system definition established by Regulation (EU) 2024/1689 (AI Act). Available at: <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application#:~:text=By%20issuing%20guidelines%20on%20the%20AI%20system%20definition%2C,on%20the%20AI%20system%20definition%20are%20not%20binding>
- European Commission, Union Rolling Work Programme - EU Cybersecurity Certification, Available at: <https://digital-strategy.ec.europa.eu/en/library/union-rolling-work-programme-european-cybersecurity-certification> .
- European Commission, New legislative framework, Available at: [https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework\\_en](https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework_en)
- European Commission, Data Governance Act explained, Available at: <https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained>
- European Commission, presentation EUHPP webinar on European Health Data Space (2/3): Secondary Use of health data (27 February 2025) of EC. Available at:

[https://health.ec.europa.eu/latest-updates/recording-and-presentation-euhpp-webinar-european-health-data-space-23-secondary-use-health-data-27-2025-03-03\\_en](https://health.ec.europa.eu/latest-updates/recording-and-presentation-euhpp-webinar-european-health-data-space-23-secondary-use-health-data-27-2025-03-03_en)

- European Data Protection Supervisor, Tech Champion: Robert Rieman, publication on ‘Synthetic Data’. Available at: [https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data\\_en](https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en)
- High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI (2019). Available at: <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>
- ICO, AI and data protection risk toolkit. Available at: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/ai-and-data-protection-risk-toolkit/>
- MDCG 2019-11 Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR (October 2019). Available at: [https://health.ec.europa.eu/system/files/2020-09/md\\_mdcg\\_2019\\_11\\_guidance\\_qualification\\_classification\\_software\\_en\\_0.pdf](https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2019_11_guidance_qualification_classification_software_en_0.pdf)

### Publications

- Alexandre Lodie, ‘Are personal data always personal? Case T-557/20 SRB v. EDPS or when the qualification of data depends on who holds them’, (November 2023). Available at: <https://hal.science/hal-04292464/document>
- Ciara Staunton and others, ‘Appropriate Safeguards and Art. 89 of the GDPR: Considerations for Biobank, Databank and Genetic Research’ (2022) 13 *Frontiers in Genetics*
- Colin Mitchell and Elizabeth Redrup Hill, ‘Are synthetic health data ‘personal data’?’. Available at: <https://www.phgfoundation.org/report/are-synthetic-health-data-personal-data#:~:text=We%20found%20that%20regulators%20and,been%20reduced%20to%20remote%20levels>
- European Data Protection Supervisor, Tech Champion: Robert Rieman, publication on ‘Synthetic Data’. Accessible at: [https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data\\_en](https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en)
- K. El Emam, L. Mosquera, and R. Hoptroff, ‘Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data’. O’Reilly Media Inc, (May 2020). Available at: [https://cdn.ttgtmedia.com/rms/pdf/Practical\\_Synthetic\\_Data\\_Generation.pdf](https://cdn.ttgtmedia.com/rms/pdf/Practical_Synthetic_Data_Generation.pdf)
- Gal, M. S., & Lynskey, O, ‘Synthetic Data: Legal Implications of the Data-Generation Revolution’, 109 *Iowa Law Review*, Forthcoming, LSE Legal Studies Working Paper No. 6/2023, (January 2023). Accessible at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4414385](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4414385) Ganey, Georgi, ‘When Synthetic Data Met Regulation’, arXiv preprint arXiv:2307.00359v1, (July 2023). Available at: <https://arxiv.org/pdf/2307.00359.pdf>
- Heiko Richter, Looking at the Data Governance Act and Beyond: How to Better Integrate Data Intermediaries in the Market Order for Data Sharing, *GRUR International*, Volume 72, Issue 5, May 2023, Pages 458–470, Available at: <https://doi.org/10.1093/grurint/ikad014>
- Jordon, J., Szpruch, L., Houssiau, F., Bottarelli, M., Cherubin, G., Maple, C., Cohen, S. N., & Weller, ‘Synthetic Data - what, why and how?’ (May 2022). Available at: [https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/Synthetic\\_Data\\_Survey-24.pdf](https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/Synthetic_Data_Survey-24.pdf)
- Khaled El Emam, ‘Precaution, ethics and risk: Perspectives on regulating non-identifiable data’, *IAPP*, (May 2022). Available at: <https://iapp.org/news/a/precaution-ethics-and-risk-perspectives-on-regulating-non-identifiable-data/>

- López, Cesar Augusto Fontanillo, 'On the legal nature of synthetic data', NeurIPS 2022 Workshop on Synthetic Data for Empowering ML Research, (2022). Available at: <https://openreview.net/pdf?id=M0KMBGL2yr>
- López, C. A. F, 'On synthetic data: a brief introduction for data protection law dummies', European Law Blog, (September 2022). Available at: <https://www.europeanlawblog.eu/pub/on-synthetic-data-a-brief-introduction-for-data-protection-law-dummies/release/1>
- Theresa Stadler, Bristena Oprisanu, Carmela Troncoso, 'Synthetic Data -- Anonymisation Groundhog Day', (November 2020). Available at: <https://arxiv.org/abs/2011.07018>
- Theresa Stadler, Bristena Oprisanu & Carmela Troncoso, 'Synthetic Data – Anonymisation Groundhog Day' (unpublished manuscript, January 2022). Available at: <https://arxiv.org/pdf/2011.07018.pdf>.
- Timo Kohlberger & Yuan Liu, 'Generating Diverse Synthetic Medical Image Data for Training Machine Learning Models', (February 2020). Available at: <https://blog.research.google/2020/02/generating-diverse-synthetic-medical.html?m=1>
- Weitzenboeck, Emily M., et al. "The GDPR and unstructured data: is anonymization possible?" International Data Privacy Law 12.3 (2022): 184-206

### Judgements

- CJEU Judgement of 22/11/2012, Case C-219/11, Brain Products GmbH v BioSemi VOF, Antonius Pieter Kuiper, Robert Jan Gerard Honsbeek,
- Alexander Coenraad Metting van Rijn, ECLI:EU:C:2012:742
- CJEU Judgement of 19/10/2016, Case C-582/14, Patrick Breyer v Bundesrepublik Deutschland, ECLI:EU:C:2016:779
- CJEU Judgement of 5/06/2018, Case-210/16 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie, ECLI:EU:C:2018:388
- CJEU Judgement of 10/07/2018, Case C-25/17, Tietosuoja-valtuutettu v Jehovan todistajat — uskonnollinen yhdyskunta, ECLI:EU:C:2018:551
- CJEU Judgement of 29/07/2019, Case C-40/17, Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW eV, ECLI:EU:C:2019:629
- General Court (Eighth Chamber, Extended Composition), Judgment of 26/04/2023, Case T-557/20, Single Resolution Board (SRB) v European Data Protection Supervisor (EDPS), ECLI:EU:T:2023:219
- Opinion of Advocate General Spielmann delivered on 6/02/2025, Case C-413/23 P European Data Protection Supervisor v Single Resolution Board, Available at <https://curia.europa.eu/juris/document/document.jsf?text=&docid=295078&pageIndex=0&doclang=EN>

### Other sources

- FUTURE-AI checklist, Available at <https://future-ai.eu/checklist/>
- CRA guide for software developers - Cyber Resilience Act, <https://www.cyberresilienceact.eu/cra-guide-for-software-developers/> CRA guide for software developers - Cyber Resilience Act, Available at: <https://www.cyberresilienceact.eu/cra-guide-for-software-developers/>
- EU Cybersecurity Certification, Available at: [https://certification.enisa.europa.eu/index\\_en](https://certification.enisa.europa.eu/index_en)
- The CRA, explained - Cyber Resilience Act, Available at: <https://www.cyberresilienceact.eu/the-cra-explained/>