Building governance for health data spaces and infrastructures: interplay between GDPR, AIA, EHDS

Ricard Martínez martinez

Director of the Chair for Privacy and Digital Transformation Microsoft-University of Valencia

Our Goals

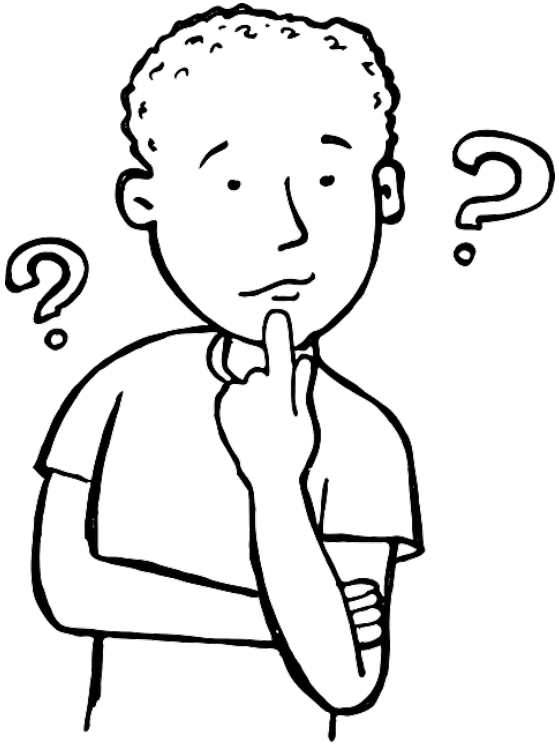
- Conditions defined by the Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space for the development of a secure health data space.
- Different data spaces configuration alternatives.
- The identification of some risks related with health data spaces from a legal point of view.

1. EHDS secure processing environment requirements

In order for a secure processing environment to be deemed to exist, Article 50 of the future regulation requires the following security measures:

(a) restrict access to the secure processing environment to authorised natural persons listed in the respective data permit;
(b) 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 technical and organisational measures;
(c) limit the input of electronic health data and the inspection, modification or deletion of electronic health data hosted in the secure processing environment to a limited number of authorised identifiable individuals;
(d) ensure that health 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;
(e) keep identifiable logs of access to and activities in the secure processing environment for the period of time necessary to verify and audit all processing operations in that environment. Logs of access should be kept for not shorter than one year;
(f) ensure compliance and monitor the security measures referred to in this Article to mitigate potential security threats.

In practice

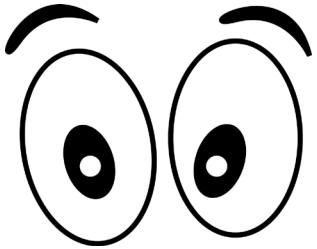


- Authentication.
- Prevent intrusions and unauthorised use.
- Clear definition of profiles and permissions granted to the user with a clear distinction between user and administrator profiles.
- Traceability
- Periodic checks and audits.

But…

- The Commission shall, by means of implementing acts, provide for the technical, organisational, information security, confidentiality, data protection and interoperability requirements for the secure processing environments, including the technical characteristics and tools available to the health data user within the secure processing environment.

Assumptions



- It is not possible to include a comprehensive description of the measures in a Regulation.
- There are national standards to be considered such as the Spanish National Security Standard.
- It is reasonable to assume that in many Member States healthcare organisations are applying some certifiable standard such as ISO 27001.
- Under the EHDS, data must be processed on European territory: this will affect external services providers.
- The Commission should define the conditions for a SPE secure data space and the interactions between :
  - Healthdata@EU infrastructure,
  - national data spaces,
  - authorised participants and
  - trusted data holders.

2. Different data spaces configuration alternatives

- Data spaces provided by the data holder on its own premises or in the cloud. This should allow it to:
- (i) Participate in federated processing environments.
- (ii) Become recognised as a trusted data holder.

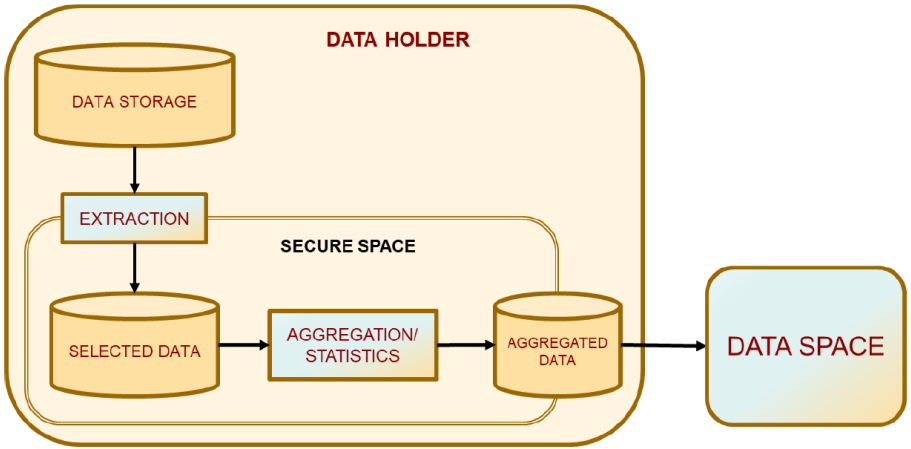


Figure 18: Diagram of the architecture for the use case of anonymised data without linkage between the data from different Data Holders

- Secure data spaces other than national data spaces:
  - (i) Trans-European projects with their own data space
  - (ii) Data altruism organisations data space
  - (iii) Any type of aggregation similar to a Marketplace
  - (iv) European infrastructures and/or authorised participants (as EUCAIM)
- National data spaces:
  - (i) Under the control of a regional or national public authority
  - (ii) Under the control of the Health Data Access Body
- HealthData@EU Space

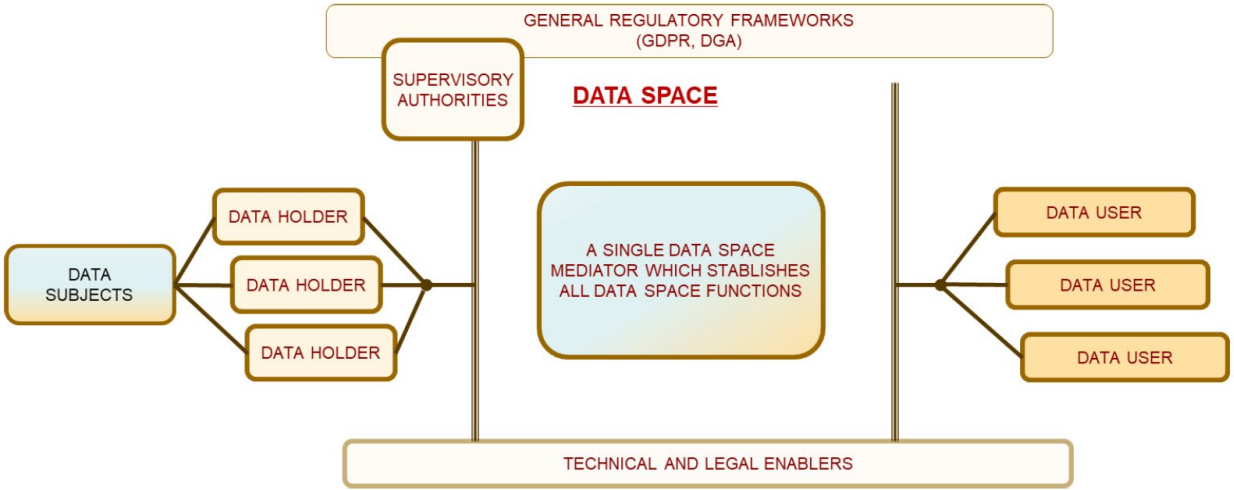


Figure 3: Configuration of a Data Space based on sharing via a central node

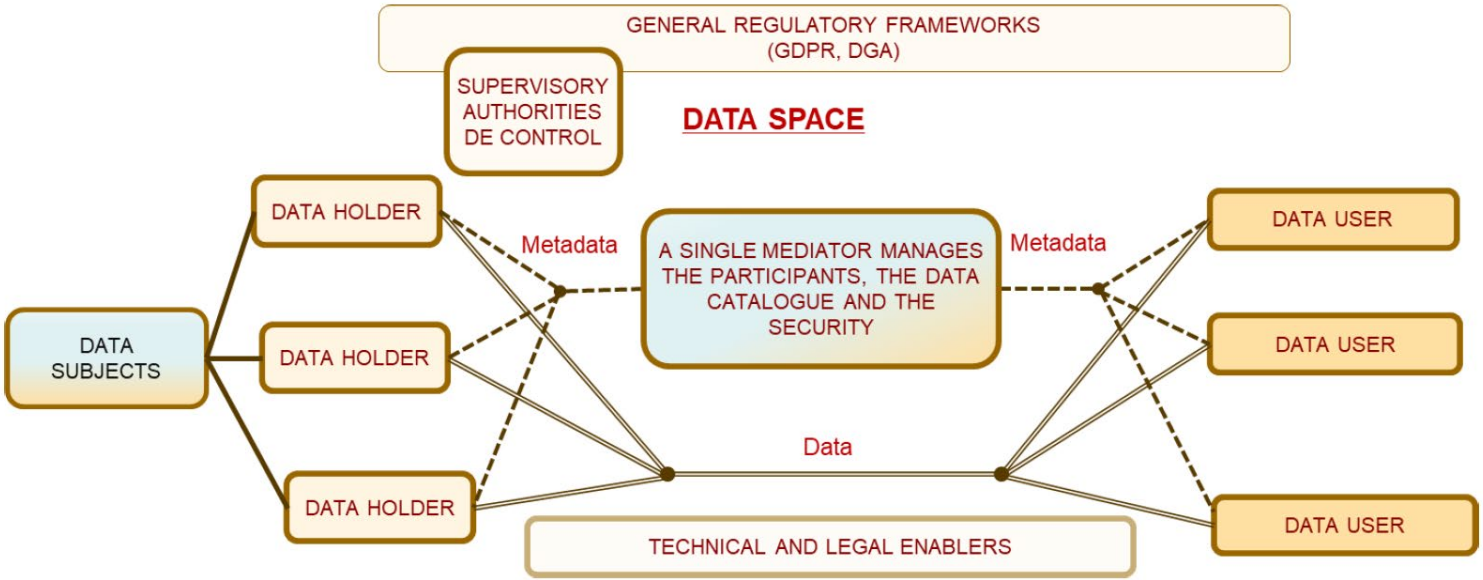


Figure 4: Configuration on the basis of a Data Space Mediator as a central hub or data marketplace

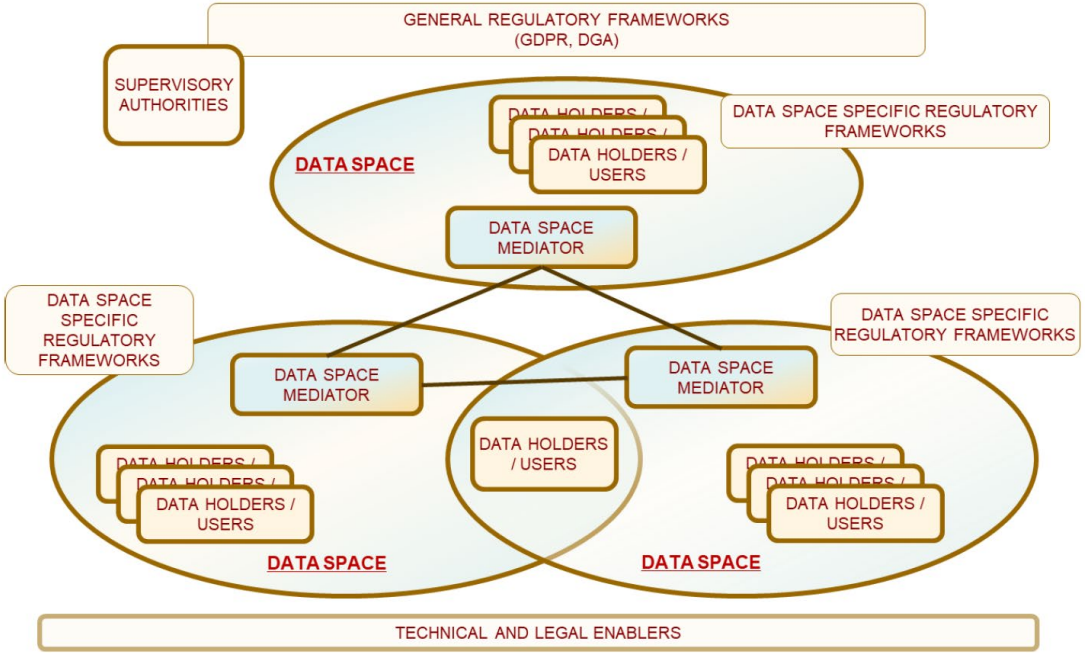
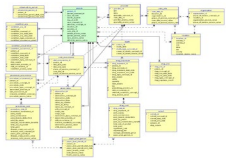


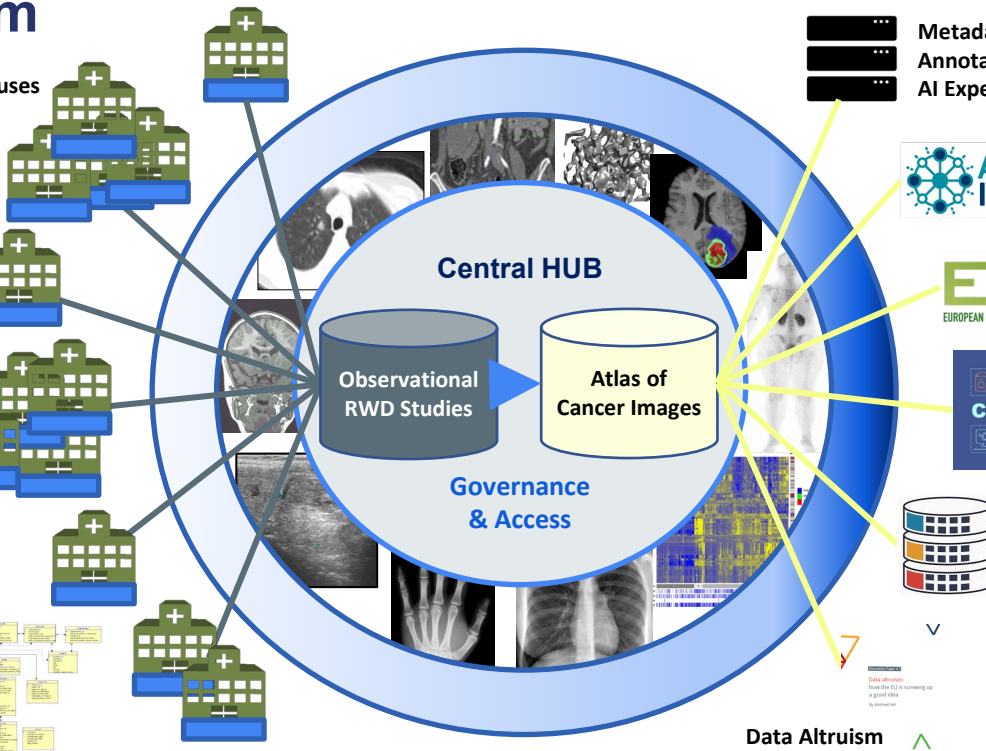
Figure 9: Federation of Data Spaces

Hybrid Platform

Distributed RW Data Warehouses
Extract, Transform and Load
ML Federated Learning

CDM hyper-ontology





Metadata Catalogue
Annotated Structured Data
AI Experimentation Platform

AI4Health Imaging

ERIC
EUROPEAN RESEARCH INFRASTRUCTURE CONSORTIUM

CLINICAL TRIALS

FAIR

DICOM-MIABIS

Data Altruism

3. some risks related with health data spaces from a legal point of view.

- The European Health Data Space promotes the trans-European use of data and like the Data Governance Act designs a closed and traceable SPE. This type of ecosystem is essential for the achievement of the goal of safeguarding the fundamental right to data protection and other rights such as intellectual property and trade secrets. .
  - (i) Software middleware will be very important for rights protection. Applications need to be able to make requests and receive results, while ensuring that two objectives are met: 1) rights assurance and 2) usage traceability.
  - (ii) In federated processing, availability will be crucial. In trans-European processing, it is very likely that a data user will program treatments that process data in different datasets. Under certain conditions this could work as a neural network. It will be necessary to ensure the availability of all federated nodes. If one or more of them fail, the results could be compromised in terms of data diversity and volume. In this case, there could be a risk of bias or a lack of explainability in the development of the artificial intelligence.

Players requirements

- Data holders must ensure that data sets have been created with the necessary ethical and legal requirements. In different Member States, this may involve:
  - (i) Ethical approval.
  - (ii) Demonstrating the lawfulness of the processing in accordance with Articles 6 and/or 9 of the GDPR. Also having conducted the appropriate risk analysis or privacy impact assessment.
  - (iii) Have reliable anonymisation and/or robust pseudonymisation procedures in place.
  - (iv) Comply with the requirements of the future European Health Data Space in areas such as cataloguing datasets and ensuring their quality and utility.
  - (v) Be able to ensure withdrawal of consent and/or opt-out of patients and maintain version control of the dataset.

- Data access applicants must also provide safeguards:
  - (i) They must comply with the requirements for applying for data access and comply with the obligations imposed on them by the European Health Data Space.
  - (ii) They should ensure compliance with applicable ethical requirements.
  - (iii) They should understand the opersational rules of the Data Space and accept its security obligations including non- re-identification commitments.
  - (iv) They should implement applicable legal obligations. For example:
    - Apply GDPR to their own datasets if they need to be involved in the processing.
    - Comply with biomedical research requirements.
    - Comply with the requirements of the Artificial Intelligence Regulation.
    - Comply with the requirements of the Medical Devices Regulation.

Staff

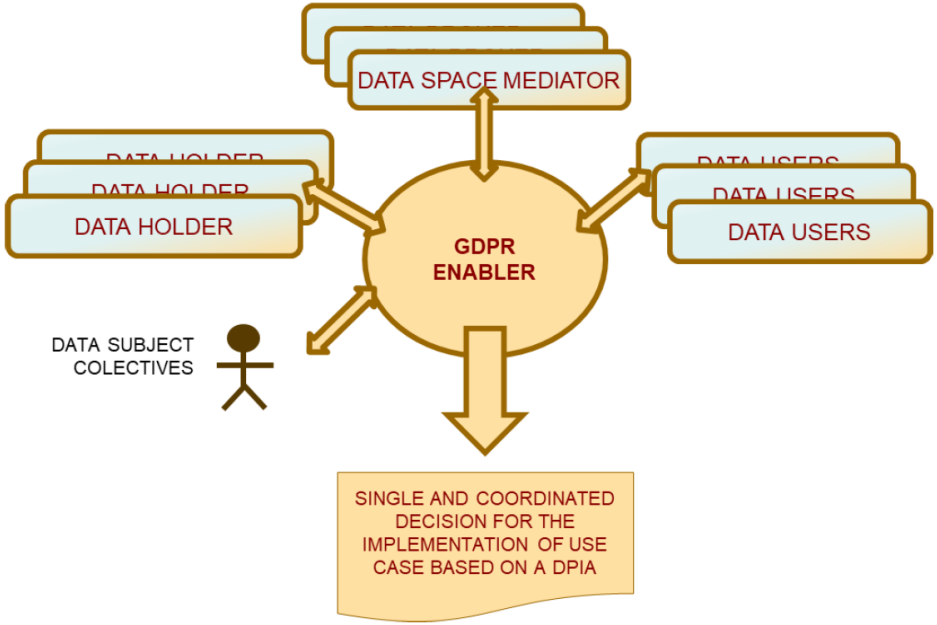
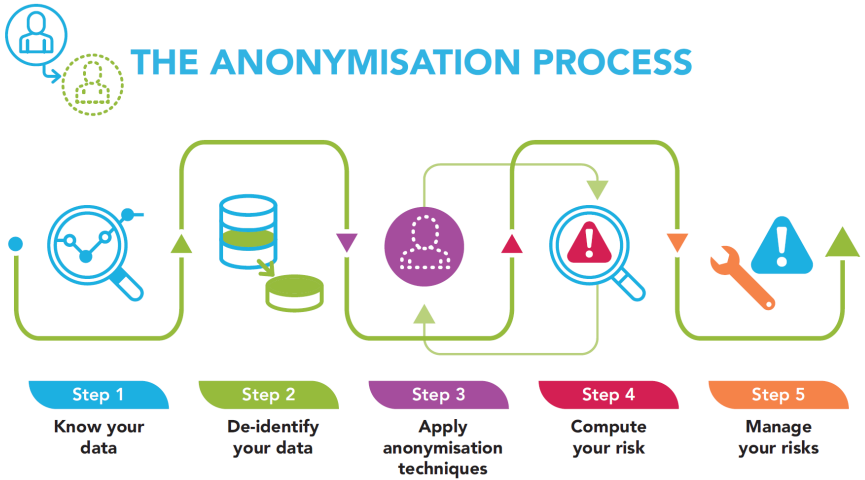


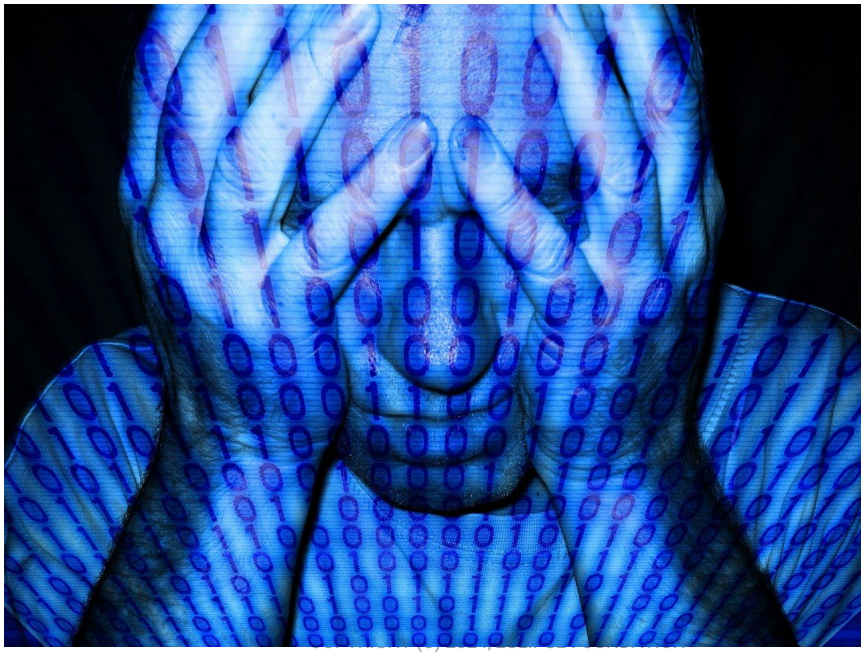
Figure 25: Role of a Data Protection Enabler for the coordination and legal, organisational and technical support to the different interveners involved in a processing operation in a Data Space

More risks?

- The DPA's conception of anonymisation is not sustainable and incompatible with the configuration of data spaces as secure processing environments. In these spaces, robust anonymisation combined with the use of techniques such as differential privacy, multi-party computation or homomorphic encryption, and traceability of use ensure adequate anonymisation.



- Informed consent requirements for each specific processing activity are technologically feasible, but emotionally impossible. In my opinion, except for patients closely involved in research, click fatigue will affect health data altruism and make them extremely difficult.



Thank you