# FAIR principles and health data, security and privacy

**Jaime Delgado**

Information Modelling and Processing Research Group
Dept. Arquitectura de Computadors
Universitat Politècnica de Catalunya (UPC)

*jaime.delgado@upc.edu*

Research Café
25th October 2023

# Presenter’s presentation

**Jaime Delgado** *40 years of research work ☺*

Main current …

- … topics: Privacy & Security on images and health/genomic information
- … projects: GenClinLab, MedSecurance (EU), Standardization (**ISO**: JPEG, MPEG, Health & Genomics Informatics, Personalized Digital Health; **IEEE**; **HL7**; …)
- … research positions: EFMI (European Federation for Medical Informatics) SEC WG Chair; Editor/ Project lead JPEG Systems RefSW, Personalized Digital Health Framework, AhG Chair JPEG Systems

# Contents

- FAIR principles
- FAIRification
- Security & Privacy
- Health data
- Conclusions

# FAIR principles

- FAIR data

- Findable
- Accessible
- Interoperable
- Reusable

# FAIR principles

## Findable

- F1: (Meta)data are assigned a globally unique and persistent **identifier**
- F2: Data are described with rich **metadata**
- F3: Metadata clearly and explicitly include the identifier of the data they describe
- F4: (Meta)data are registered or **indexed** in a searchable resource

# FAIR principles

## Accessible

- A1: Meta(data) are retrievable by each of their identifiers using a standardized communication **protocol**
  - A1.1: The protocol is **open**, free and universally implementable
  - A1.2: The protocol allows for an **authentication and authorization**, where necessary
- A2: Metadata should be accessible even when the data is **no longer** available

# FAIR principles

## Interoperable

- I1: Metadata and data use a formal, accessible, shared, and broadly applicable language for **knowledge representation**
- I2: Metadata and data use **vocabularies** that follow the FAIR principles
- I3: Metadata and data include **qualified references** to other metadata and data

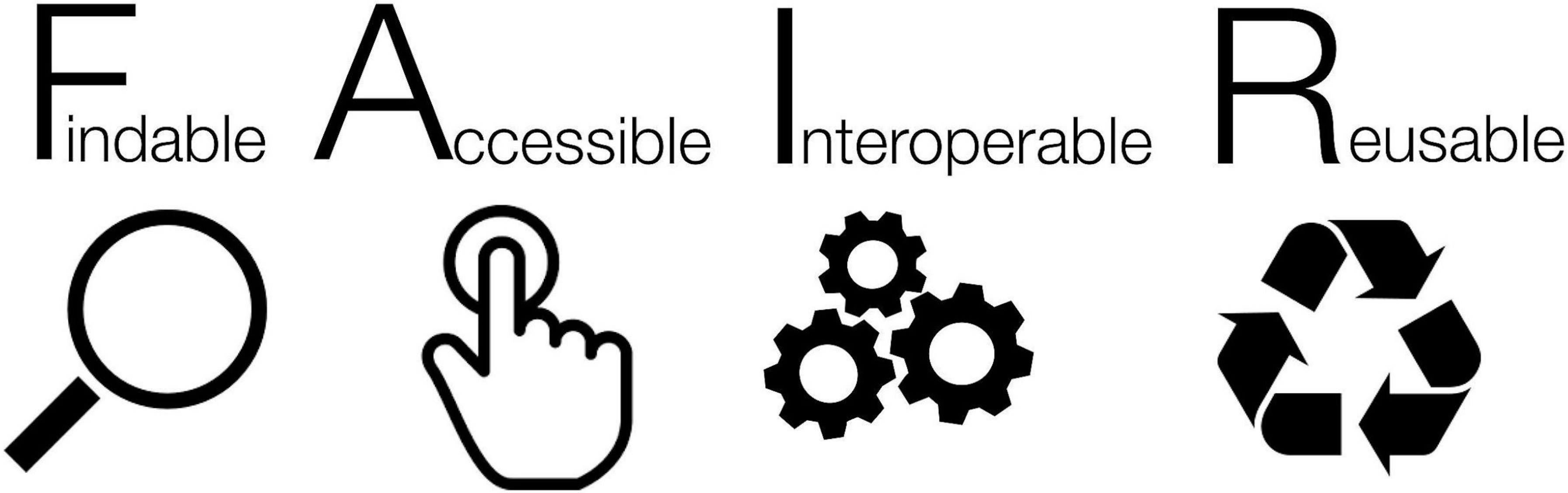
# FAIR principles

## Reusable

- R1: Metadata and data are richly described with a plurality of accurate and relevant **attributes**
  - R1.1: Metadata and data are released with a clear and accessible **data usage license**
  - R1.2: Metadata and data are associated with detailed **provenance**
  - R1.3: Metadata and data meet domain-relevant **community standards**

# FAIR principles

- FAIR data



Findable Accessible Interoperable Reusable

# FAIRification initiatives

- Guidelines to help in making the data FAIR

GO FAIR

FAIR4Health

RDA

RESEARCH DATA ALLIANCE

FORCE11

The Future of Research Communications and e-Scholarship

FAIRsharing.org

(registry of standards)

# FAIRification workflow steps (FAIR4Health)

1) Raw data analysis
2) Data curation & validation
3) Data de-identification / anonymization
4) Semantic modeling
5) Make data linkable
6) License attribution
7) Data versioning
8) (Meta)data aggregation
9) Archiving

# Contents

- FAIR principles
- FAIRification
- **Security & Privacy**
- Health data
- Conclusions

# FAIRification workflow steps (SECURITY)

1) Raw data analysis
2) Data curation & validation
3) Data de-identification / anonymization
4) Semantic modeling
5) Make data linkable
6) License attribution
7) Data versioning
8) (Meta)data aggregation
9) Archiving

# License attribution

- We focus on ***license attribution*** step
- Framework for data owners to provide licensing
- Support proper reusability (FAI**R**)
- Request permission to use (may include authentication & authorization) (F**A**IR)

- *Absence of explicit license may prevent to reuse data*

# License attribution – Problems & solutions

**We need to solve:**

1. How to express the licenses?
2. How to guarantee their provenance?
3. How to evaluate their authorization?
4. How to enforce what they are controlling?

# Expression – proposed solution

- **1. How to express licenses?**
  - Formal language □ Interoperability (FAIR)
  - Rules formally expressed □ Clearly define access to information (FAIR)
- Option: eXtensible Access Control Markup Language (XACML)

# Expression of licenses

- **XACML**
  (eXtensible Access Control Markup Language)
- Express privacy rules/policies *(OASIS standard)*
- Control
  who, how and under which conditions
  access specific information
  (data or metadata)
- Mechanism to evaluate the rules (authorize), based on standardized *requests*

# Provenance – proposed solution

- **2. How to protect provenance?**
  - Digital signature □ XML signature (F**A**IR) but also (FA**I**R) and partially (FAI**R**)

# Authorization – proposed solution

- **3. How to authorize?**
  - Using *XACML Requests* □ Access control & Interoperability (F**AI**R)
  - Attributes: subject, object (data or metadata), action, time conditions, …

# Enforcement – proposed solution

- **4. How to enforce?**
- Protect from unauthorized access (F**A**IR)

# Contents

- FAIR principles
- FAIRification
- Security & Privacy
- **Health data**
- Conclusions

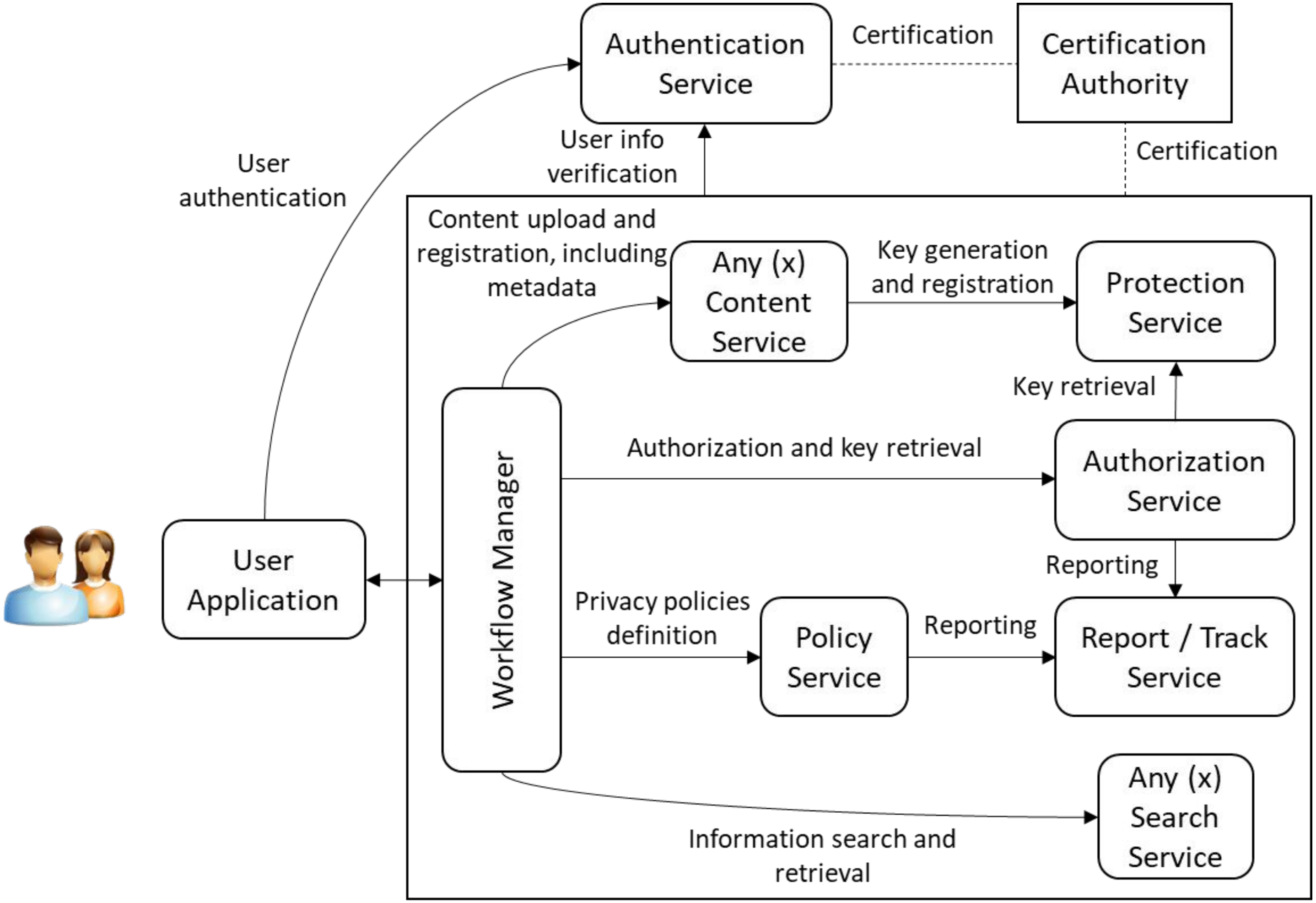
# Application to health data

- Modular and distributed approach for the management of health information □ *Health Information Protection And Management System* (HIPAMS)
- Support of FAIR principles from a security and privacy point of view
- Focus on privacy rules to control the access to information

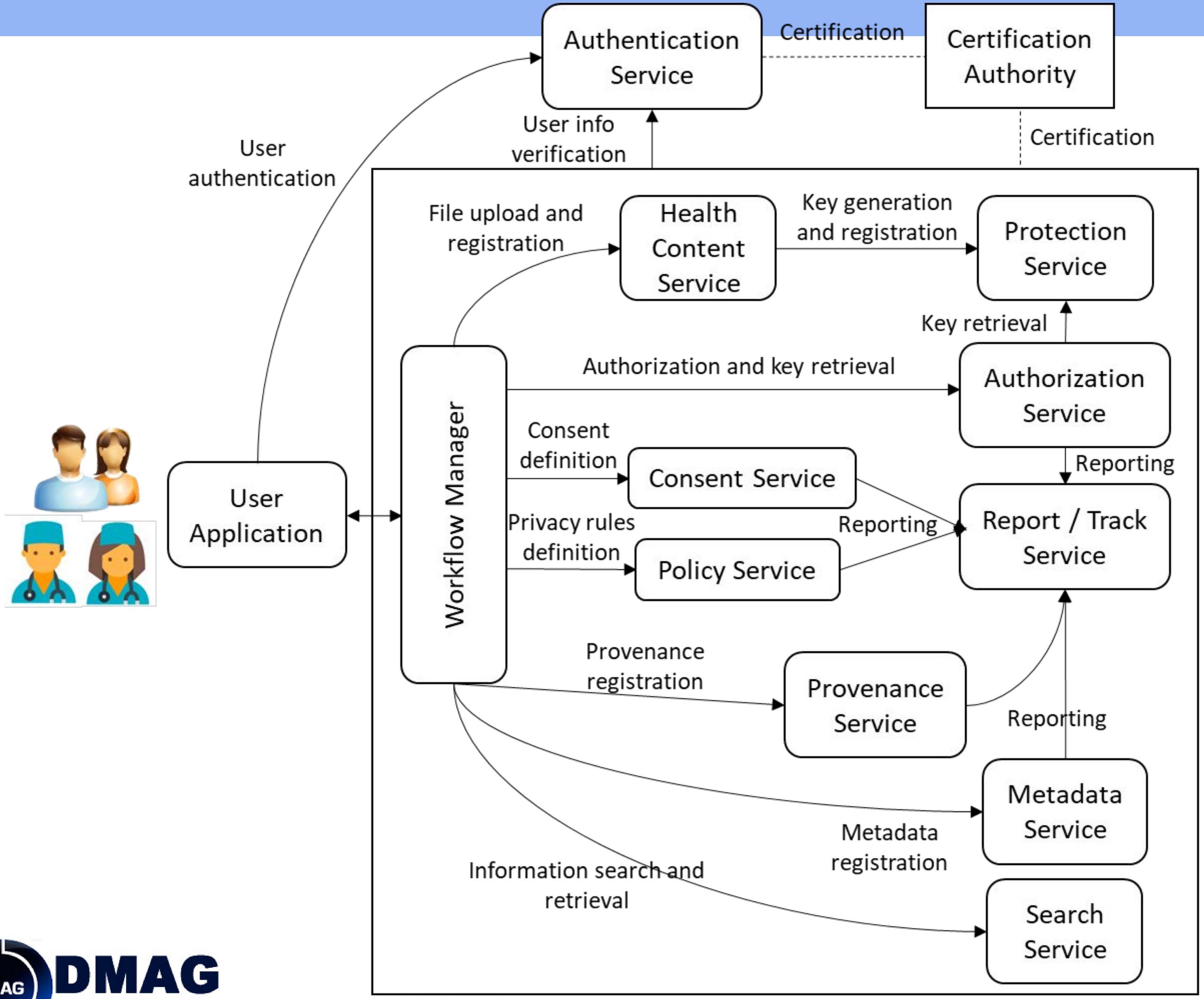
# xIPAMS approach

- Architecture independent of the kind of content:
  - Different "xIPAMS" platforms possible:
    - MIPAMS (Multimedia)
    - GIPAMS (Genomic)
    - HIPAMS (Health)
  - Possible integration of genomic and other health information
- Apply this to different projects

# xIPAMS Architecture



# HIPAMS Architecture



DMAG
DISTRIBUTED MULTIMEDIA APPLICATIONS GROUP

# *License attribution* FAIRification step – HIPAMS solution

- **1. How to express licenses?**
  - Formal language □ Interoperability (FA**I**R)
  - Rules formally expressed □ Clearly define access to information (F**A**IR)
- Option: eXtensible Access Control Markup Language (XACML)
- HIPAMS module □ Policy Service (privacy policies creation)

# *License attribution* FAIRification step – HIPAMS solution

- **2. How to protect provenance?**
  - Digital signature □ XML signature (F**A**IR) but also (FA**I**R) and partially (FAI**R**)
- HIPAMS module □ Policy Service (privacy policies creation)

# *License attribution* FAIRification step – HIPAMS solution

- **3. How to authorize?**
  - Using *XACML Requests* □ Access control & Interoperability (F**AI**R)
  - Attributes: subject, object (data or metadata), action, time conditions, …
- HIPAMS module □ Authorization Service (privacy policies authorization)

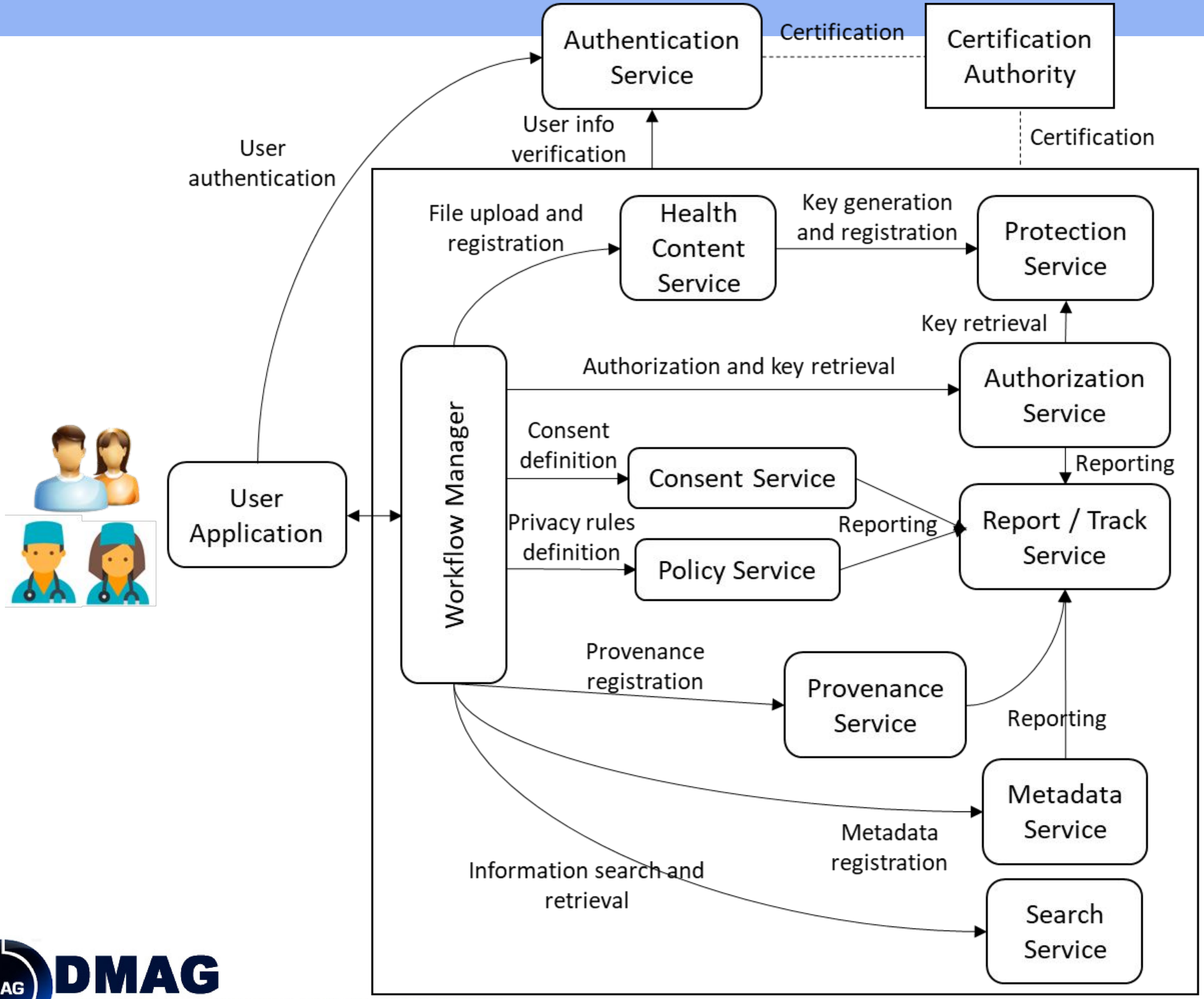
# *License attribution* FAIRification step – HIPAMS solution

- **4. How to enforce?**
- Protect from unauthorized access (F**A**IR)
- HIPAMS modules:
- Content to provide in *Health Content Service*
- Content encrypted with *Protection Service*

# *License attribution* FAIRification step – HIPAMS solution

- **Globally:**
  - HIPAMS modules □ All the platform!
- In addition:
  - Standardized formats in *Health Content Service* □ Interoperability (FA**I**R)
  - Keeping track of the actions with *Reporting Module*
  - *Search Service* □ Findability! (**F**AIR)

# HIPAMS Architecture



Authentication Service

Certification

Certification Authority

Certification

User info verification

User authentication

File upload and registration

Health Content Service

Key generation and registration

Protection Service

Key retrieval

Authorization and key retrieval

Authorization Service

Workflow Manager

User Application

Consent definition

Consent Service

Reporting

Privacy rules definition

Policy Service

Reporting

Report / Track Service

Provenance registration

Provenance Service

Reporting

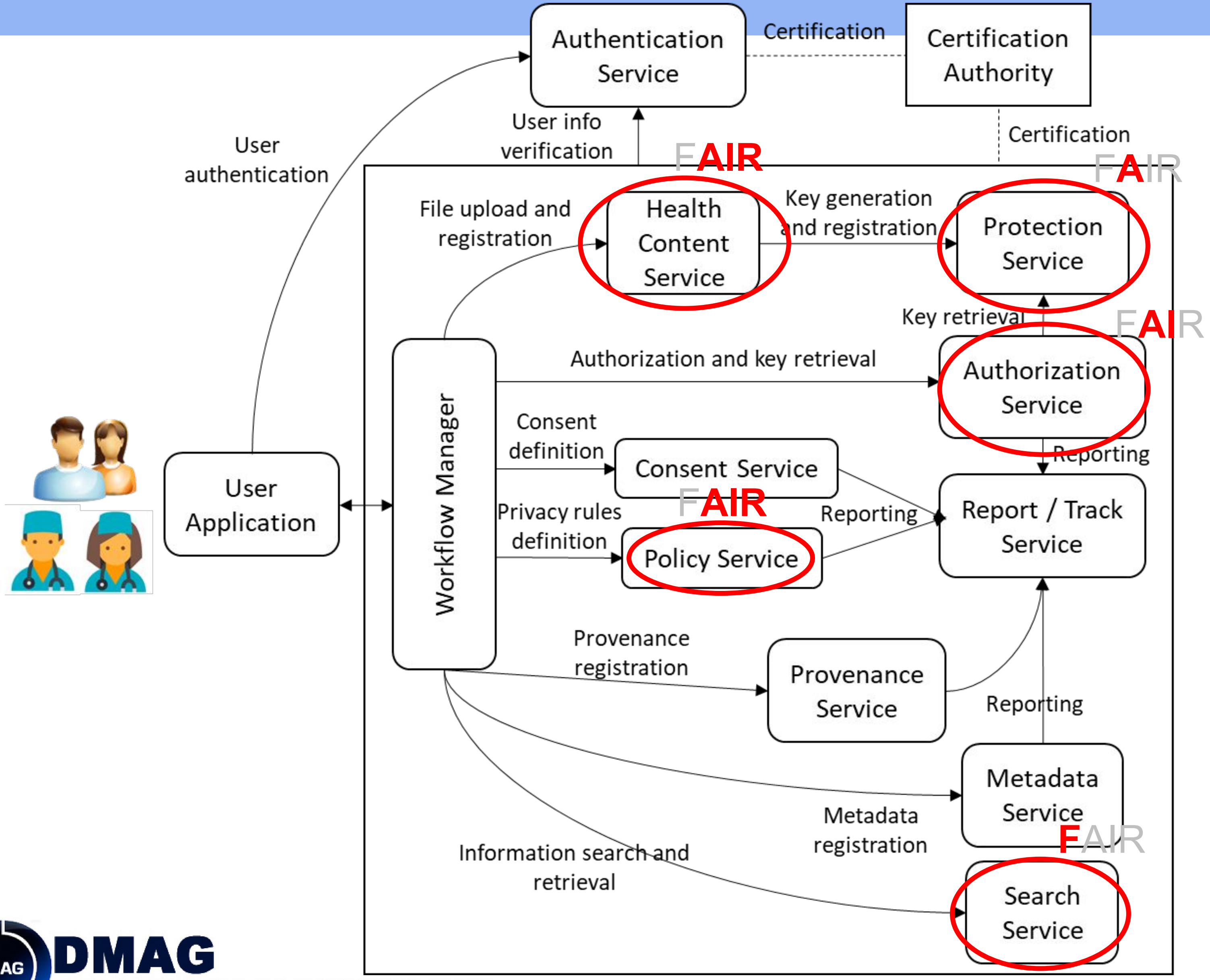
Metadata Service

Metadata registration

Information search and retrieval

Search Service

# HIPAMS Architecture



DMAG
DISTRIBUTED MULTIMEDIA APPLICATIONS GROUP

# Conclusions

- FAIR principles, basis for improving the use of existing data
- Data to be “FAIRified”
- Health data is a specific case. Access and distribution to be controlled, but open for research (privacy-aware)
- Security & Privacy mechanisms available

# FAIR principles and health data, security and privacy

**Jaime Delgado**
Information Modelling and Processing Research Group
Dept. Arquitectura de Computadors
Universitat Politècnica de Catalunya (UPC)

*jaime.delgado@upc.edu*

Research Café
25th October 2023