



HEALTHYCLOUD
Health Research & Innovation Cloud

Flexible and responsible manner: Is it necessary to mobilize data?

Laura Portell-Silva (BSC) · Eva Garcia-Alvarez (BBMRI-ERIC)



HealthyCloud project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement N° 965345.

Sharing is caring, and it is the right thing to do

Motivations

- Accelerating research
- Increased visibility and discovery of your research data
- Increased (re)use of your research data
- Compliance with funders, publishers and other policies

Challenges

- Understanding data protection
- Use of **(meta)data standards**
- No data **repository**
- Missing skills or resources



Improving data flow and secondary data use across teams and between disciplines

Increase **Flow of Information**

Support Collaboration

Increase knowledge

Team Science

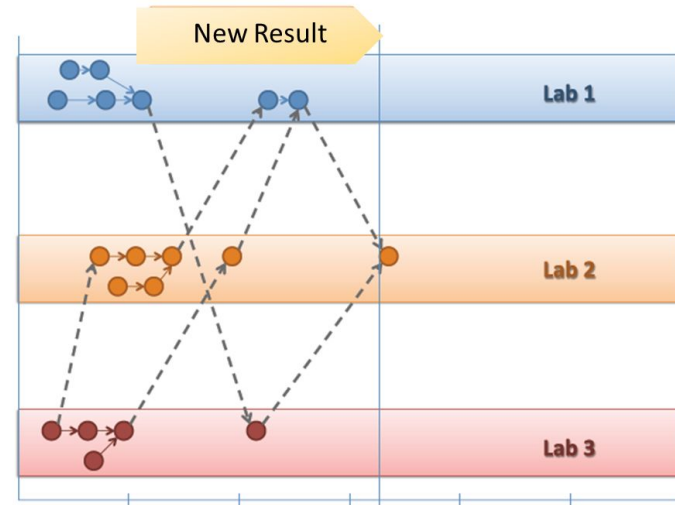
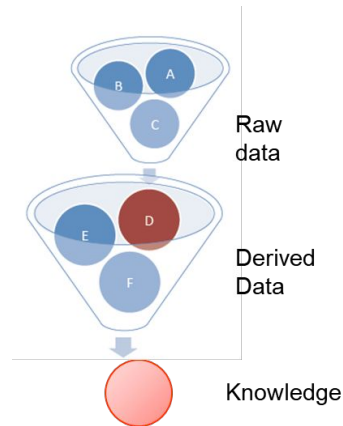


Figure: Knowledge Turning, Information Flow Josh Sommer, Chordoma Foundation, 2011

We all produce as well as reuse (other people's) data digital objects

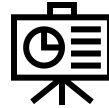


HEALTHYCLOUD
Health Research & Innovation Cloud



Documents

(Thesis, Report, Conference Papers, Articles, Survey etc.)



Slides



Datasets



Algorithms



SOPs



Codes



Figures, Photos



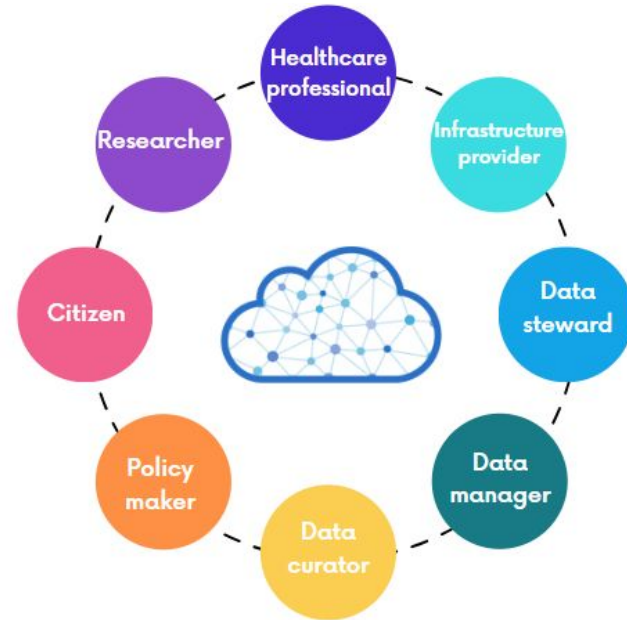
Software



Workflows



Databases



Deliverable 6.1 • “Expected user interactions”



Can data be truly shared? ... and mobilized?

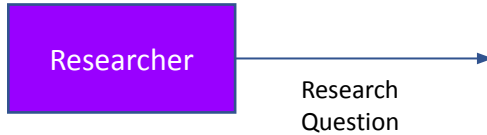


HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation

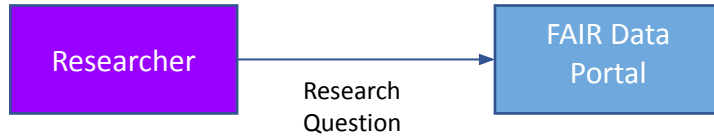


HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation



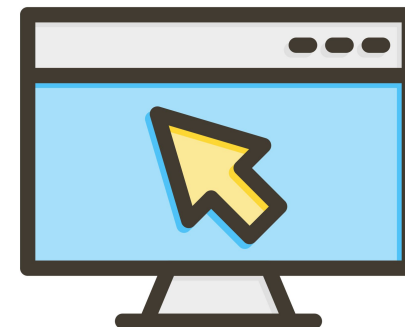
The FAIR Health Data Portal



HEALTHYCLOUD
Health Research & Innovation Cloud

FAIR health data portal specifications

- Metadata catalogue of data hubs and data collections
- Metadata contribution guidelines
- Data access conditions
- Infrastructure providers for computational resources
- Guidance & knowledge hub



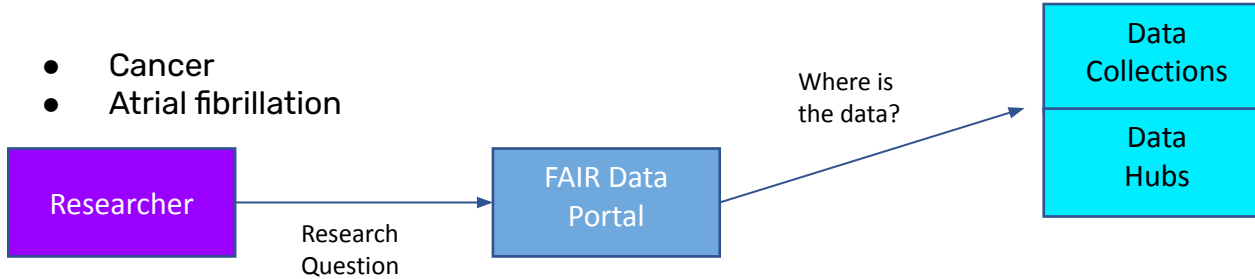
Deliverable 6.2 · “Specifications for the FAIR data portal”

HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation



Specifications of data access



Data access scenarios, from very basic functionality to most advanced interaction (full provision)

- Considerations for data access (landscape)
 - Data characteristics
 - Organization of the data sources
 - Traditional vs machine-driven data access
- Identification of data access main steps and their characteristics based on the different scenarios
 - Data access application
 - Data access negotiation
 - Data access conditions
 - Implementation of data access

Specifications need to be flexible and are highly dependent on the level of complexity of the portal

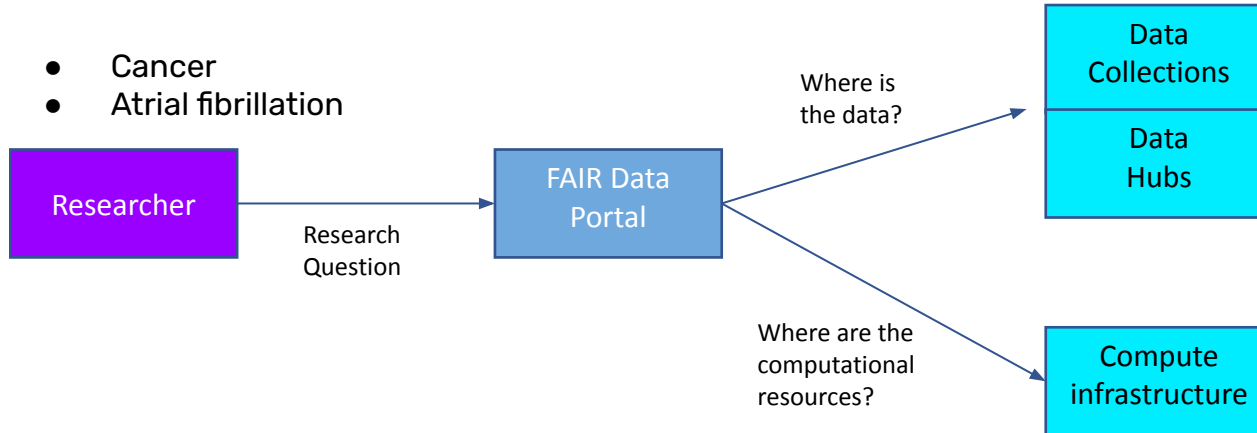
Deliverable 6.3 • “Specifications for data access”

HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation



Compute infrastructure considerations

Can data be mobilized?



HEALTHYCLOUD
Health Research & Innovation Cloud

Do I have enough computing capacity where I have the data?

- If YES, you are ready to analyse your data.
- If NOT consider using a **Secure Processing Environment (SPE)**. If data cannot be moved (GDPR, confidentiality), the SPE has to be located in the same place (depending on national and institutional agreements) where the data is. If data can be moved (anonymized, public), a global SPE can be used.

Deliverable 5.3 • “Guidelines to establish sustainable computational infrastructures for the future HRIC ecosystem”

Compute infrastructure considerations

Can data be mobilized?



HEALTHYCLOUD
Health Research & Innovation Cloud

Do I need to use data from someone else?

- If computation is available and **data cannot be moved** → Federated analysis
- If computation is available elsewhere and **data can be moved** → Distributed analysis

[Deliverable 5.2](#) • “[Analysis of existing orchestration mechanisms for distributed computational analyses including a general overview to facilitate new developments](#)”

Compute infrastructure considerations

A matter of trust...



HEALTHYCLOUD
Health Research & Innovation Cloud

- **Safe projects:** is this use of the data appropriate?
- **Safe people:** can the research team be trusted to use this data appropriately?
- **Safe data:** is there a disclosure risk in the data itself?
- **Safe settings:** does the facility itself limit unauthorised access and use?
- **Safe outputs:** are all results removed from the safe setting non-disclosive?

Deliverable 5.1 • “Analysis of existing computational infrastructures”

Compute infrastructure considerations

Enabling health research



HEALTHYCLOUD
Health Research & Innovation Cloud

- Define a capabilities maturity model for establishing different levels of engagement.
- Consider both perspectives: service providers and end-users, to facilitate expectation management.
- Look for alignment with parallel efforts, e.g. Gaia-X and EOSC.

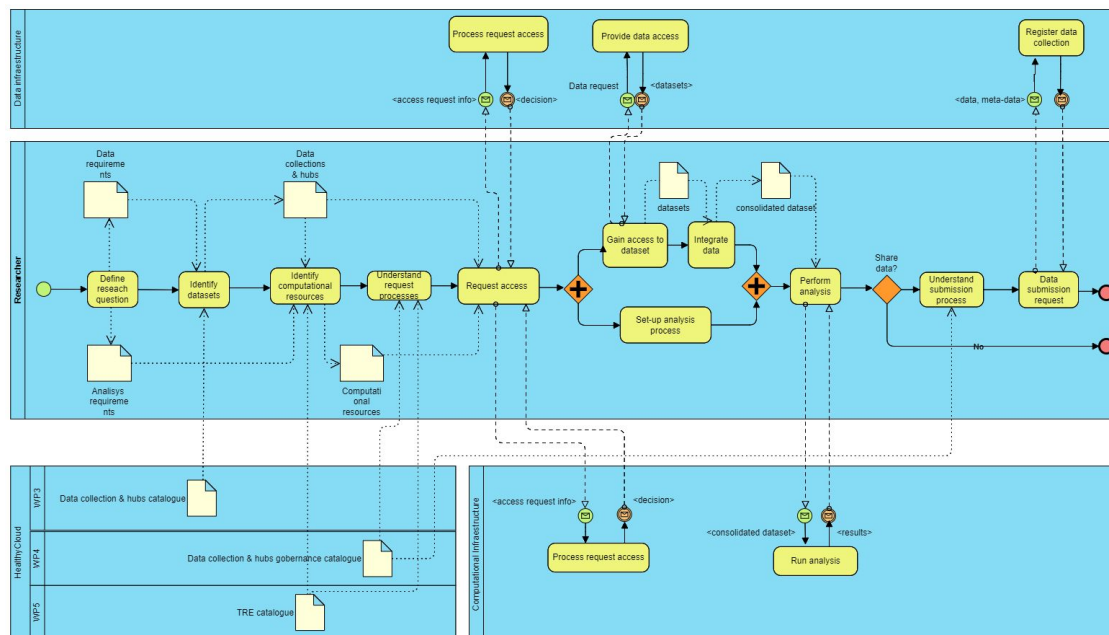


Deliverable 5.5 • “Reference guidelines for the establishment of an ethically sound and legal compliant health data research ecosystem”

HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud



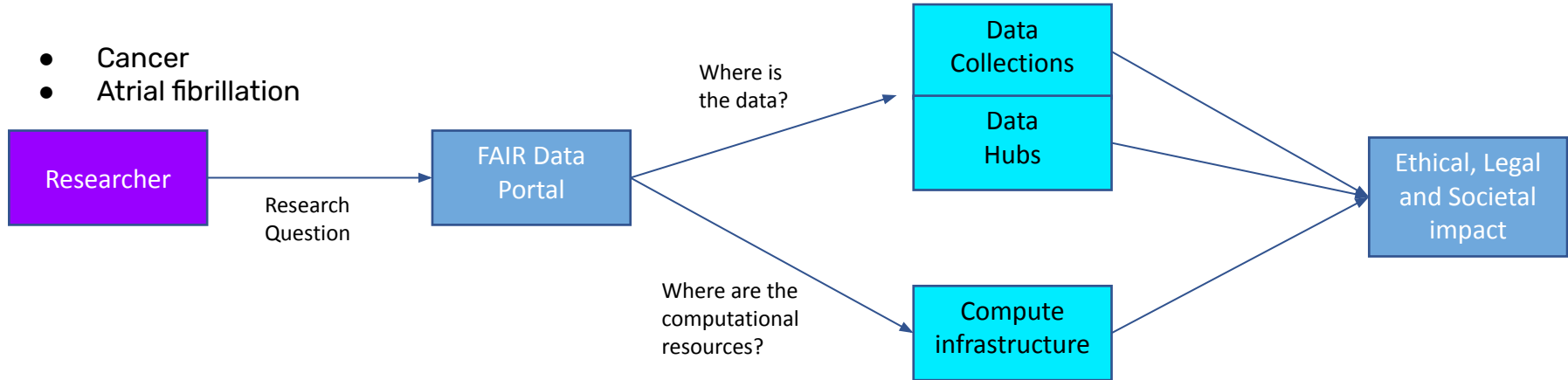
Deliverable 7.3 • “Validation of proposed solutions and gap analysis”

HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation



Conclusions



HEALTHYCLOUD
Health Research & Innovation Cloud

- The **user journey** serves as a framework for identifying key stages for **systematically** using health-related data for research purposes across Europe.
- The **FAIR Health Data Portal** is an invaluable starting point for finding available both research data and computational resources.
- **Data access conditions** play an important role as they represent the implementation of the **data governance**.
- **Data processing** is conditioned by various factors, including computational capabilities and data granularity.
- **Collaborative alignment** with other initiatives is essential to ensure a coherent framework where researchers can use health-related data seamless.

Thank you!

Follow us on social media:

www.healthycloud.eu

 @HealthyCloudEU

 www.linkedin.com/company/healthycloud.eu



HEALTHYCLOUD
Health Research & Innovation Cloud



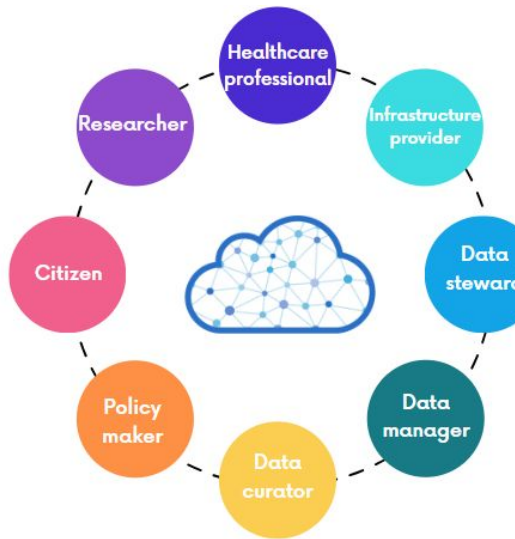
The HealthyCloud project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement N° 965345.



HEALTHYCLOUD
Health Research & Innovation Cloud

Highlights

Expected users' interactions



MARIA
CITIZEN

Age: 25
Occupation: Science teacher
Interests: Research

PHILIP
POLICY AND
DECISION MAKER

Age: 50
Occupation: Policy officer
Interests: Politics

MARIA
CITIZEN

Age: 25
Occupation: Science teacher
Interests: Research

GOALS

- Want to share their data for research, innovation and policy making
- Want to know what research is conducted for their medical condition and how the data is used
- Want to have access to summarized information about healthcare trends in the general and/or disease-specific population

NEEDS

- Find what research is being done for their condition/disease
- Where their data is being used and how
- Understand how secure the ecosystem is (data and communication)
- Know who is accessing the data (researchers or companies)

CHALLENGES

- Do not know disease
- Do not know information and policies
- Do not know

EXPECTATION

- The whole system
- Easy-to-use

PHILIP
POLICY AND
DECISION MAKER

Age: 50
Occupation: Policy officer
Interests: Politics

SKILLS

- Know how to combine information from different sources to promote new policies and adjust existing ones
- Understand the challenges for a legit use of health-related data for research purposes.

GOALS

- Want to promote policies enabling health-related data research in an ethically sound, legally compliant and socially aware manner
- Want to understand the up-take of current policies and existing barriers preventing their implementation
- Want to promote best practices for translating data-driven research results into good clinical practice guidelines
- Want to identify and extract information about general and/or disease-specific population trends to implement healthcare plans
- Want to increase the impact of the policies they are promoting
- Want to have access to aggregated information for data usage patterns from different sources and/or domains

CHALLENGES

- Do not know disease
- Do not know information and policies
- Do not know

EXPECTATION

- The whole system
- Easy-to-use

"I want to advance on policies for a trustworthy sensitive data usage"

NEEDS

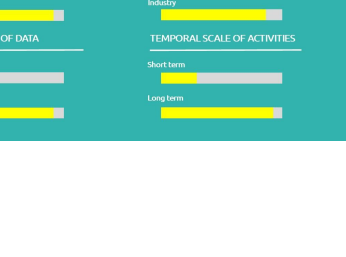
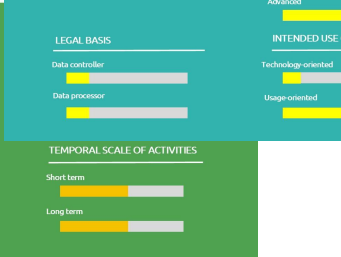
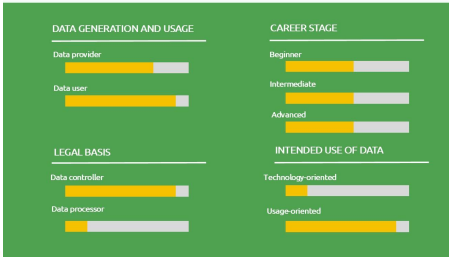
- Access to summarized information about health-related trends in the general and/or disease-specific population
- A reference place to gain access to heterogeneous health-related data sources, including aggregated information about specific healthcare aspects or data usage patterns
- Measurement of impact of the policies they put forward

CHALLENGES

- Do not know where to find health-related general population or disease-specific trends
- Access to heterogeneous data sources, which might be geographically distributed and may fall under different legal frameworks
- Justify investments and health-related decisions to tax-payers and voters

EXPECTATIONS

- Easy-to-filter trends and visualizations to enable the policy formulation and decision-making process
- A portfolio of indicators for measuring health-related data usage impact



Contribution to HRIC



HEALTHYCLOUD
Health Research & Innovation Cloud

- **Service 2:** Identification of gaps for data governance
- **Service 3:** Identification of gaps for data interoperability
- **Service 4:** A health research community interface service, with the EOSC

Lessons learned (1)



- What would you have done differently?
 - **A broader participants expertise in the workshops**
 - **In-person workshop/s to validate the findings on the user profiles**
 - **Better definition of data and computational resources access**
 - **Better definition of access conditions of the expected for the FAIR health data portal**

Lessons learned (2)



HEALTHYCLOUD
Health Research & Innovation Cloud

- Which have been the missing discussions?
 - **Collaboration with data related work packages to elaborate on services required by end-users, e.g. data anonymization, adoption of the FAIR principles, data quality assessment.**
 - **Discussions about the accessibility of computational resources**

Future directions



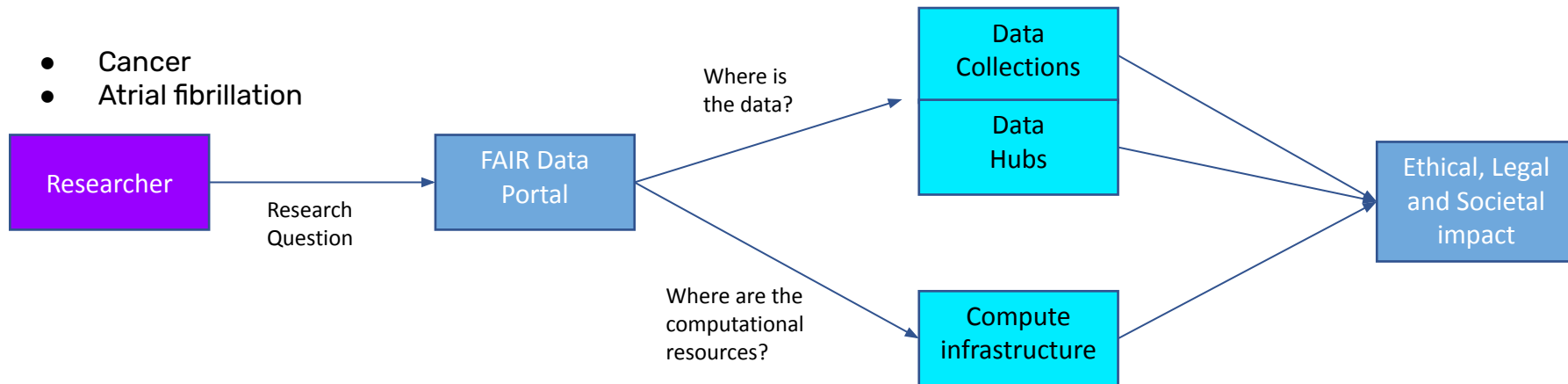
HEALTHYCLOUD
Health Research & Innovation Cloud

- What is the missing expertise in the consortium?
 - **User experience and engagement experts**
 - **End-users of the HRIC**
 - **Broader EOSC and EHDS representatives**

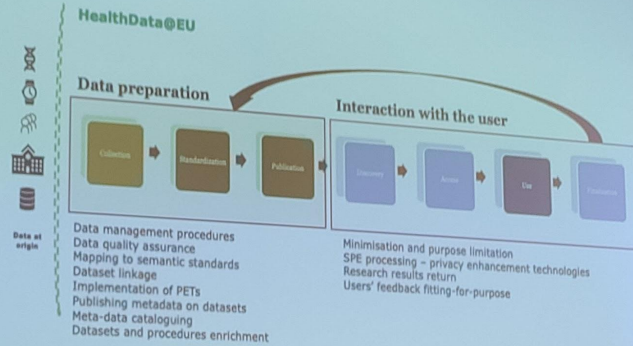
Use cases scenario



- Cancer
- Atrial fibrillation



Data life-cycle



Key recommendation: Include data quality measures as a mandatory part of the data life-cycle

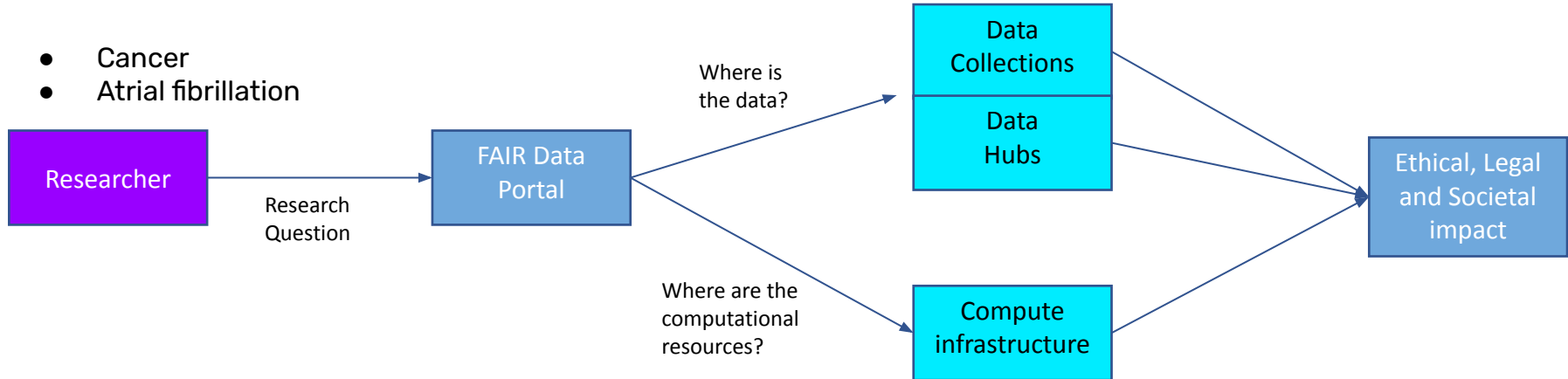
Include the V
logo in the Master Slide

HealthyCloud user story



HEALTHYCLOUD
Health Research & Innovation Cloud

- Cancer
- Atrial fibrillation



Uses cases driven analysis

