European Genomic Data Infrastructure

GDI: Genomic Data Infrastructure

Marco J. Morelli

Head of Bionformatics

Center for Omics Sciences

San Raffaele Scientific Institute

1+MG Declaration of cooperation - April 2018

https://ec.europa.eu/digital-single-market/en/european-1-million-genomes-initiative

24 countries have now signed; 4 are observers

Declaration of Cooperation (2018)
Drive personalised medicine & health in Europe

- **Vision 1+ million genomes**
  - Citizens' needs are at the centre of **data-driven healthcare innovation** as active agents in their own health journey
  - Better **prevention**, early **diagnosis** and **treatment** of diseases
  - EU stays at **the forefront of genomic and personalised medicine** globally and improves **scientific capabilities** and **industrial competitiveness**

- **One of the key goals to realise this vision**
  - "Build a **cohort of at least 1 million sequenced genomes** that provides **clinically impactful** associations in research", hence "1+MG"

Implementation projects

The vision

Cross-border access to genomic data, implementation of genomics-based health

European Commission
Horizon 2020
European Union funding
for Research & Innovation

B 1MG Beyond one Million Genomes

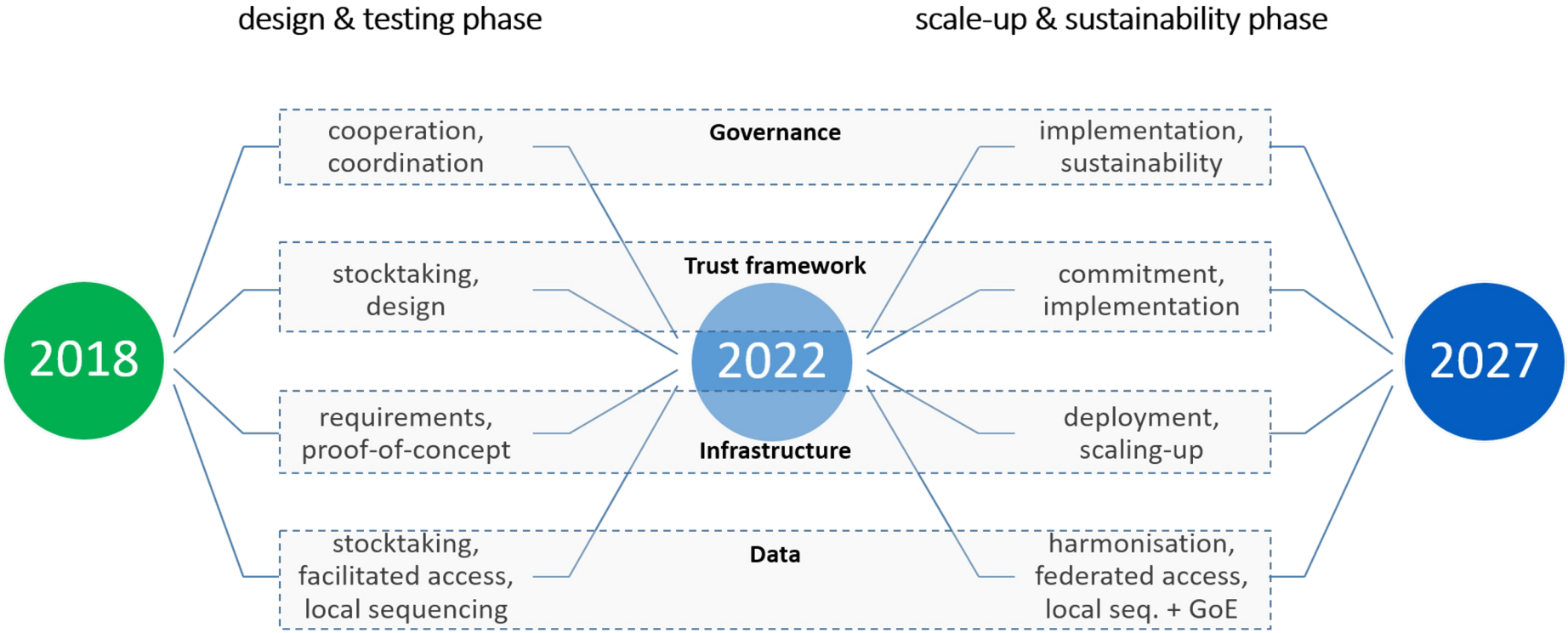
European Genomic Data Infrastructure

DIGITAL EUROPE PROGRAMME

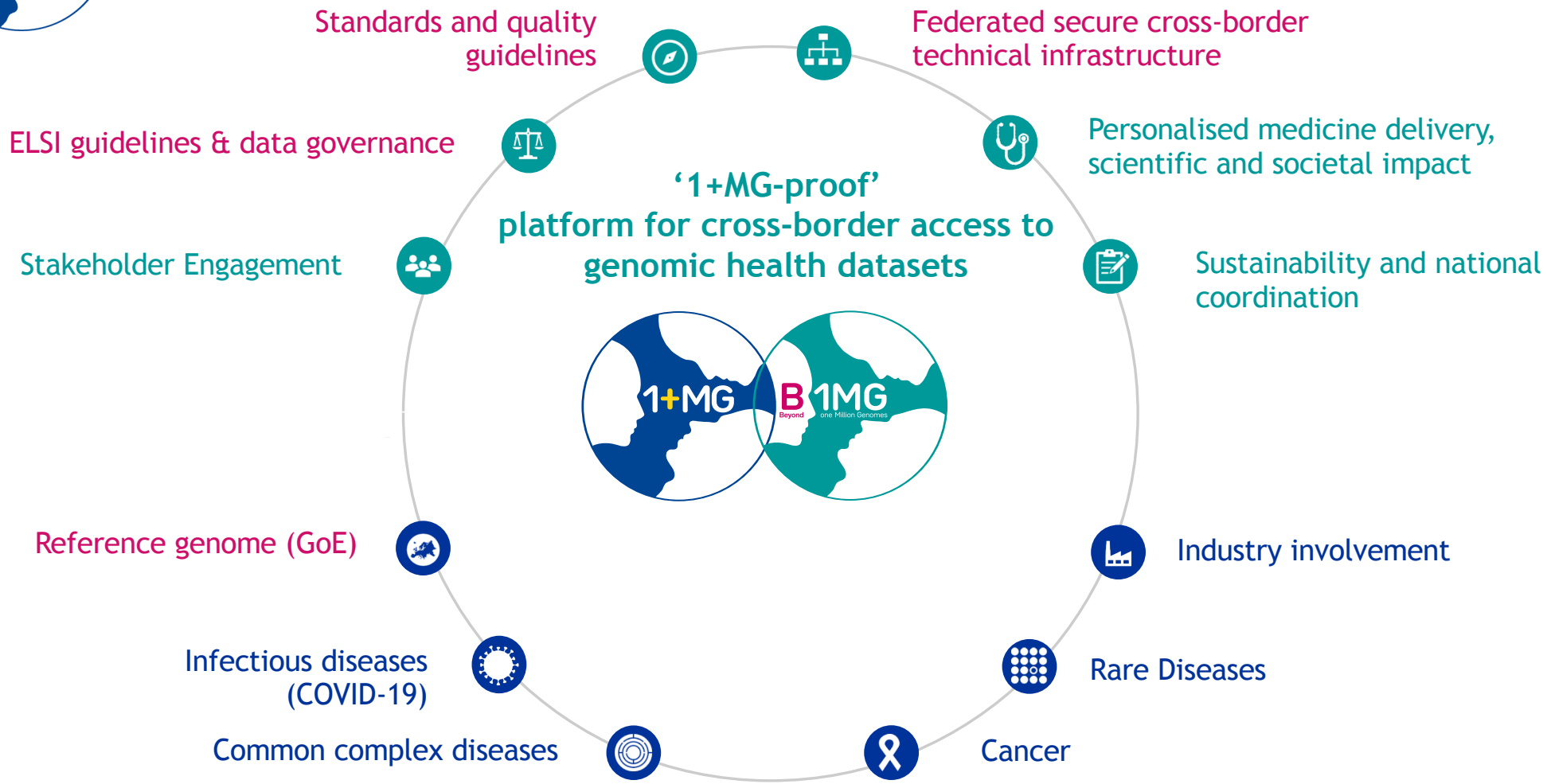
Designing and Testing

Scaling up and Sustaining

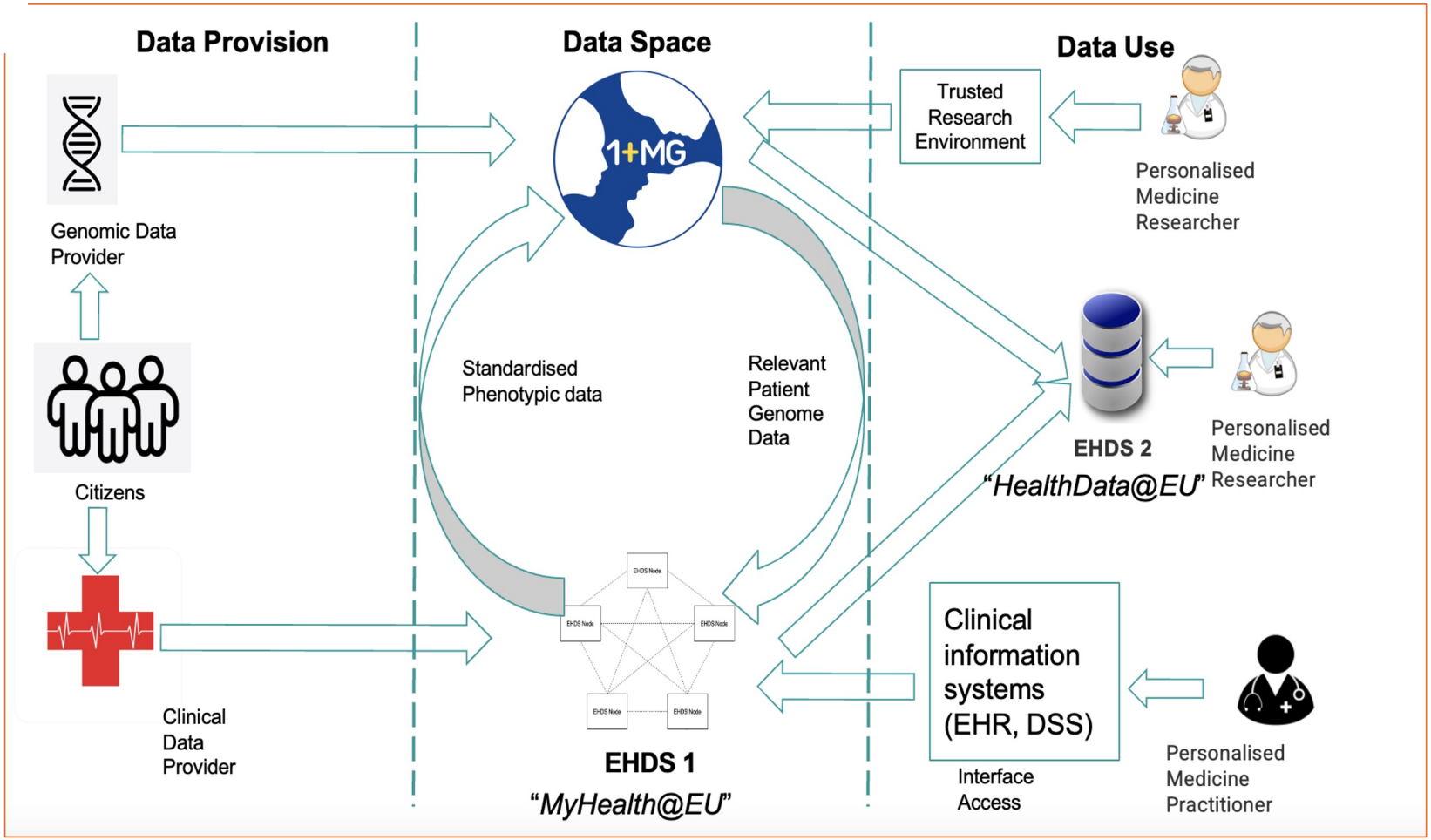
1+MG Roadmap


design & testing phase
scale-up & sustainability phase
2018
2022
2027
Governance
cooperation, coordination
implementation, sustainability
Trust framework
stocktaking, design
commitment, implementation
Infrastructure
requirements, proof-of-concept
deployment, scaling-up
Data
stocktaking, facilitated access, local sequencing
harmonisation, federated access, local seq. + GoE

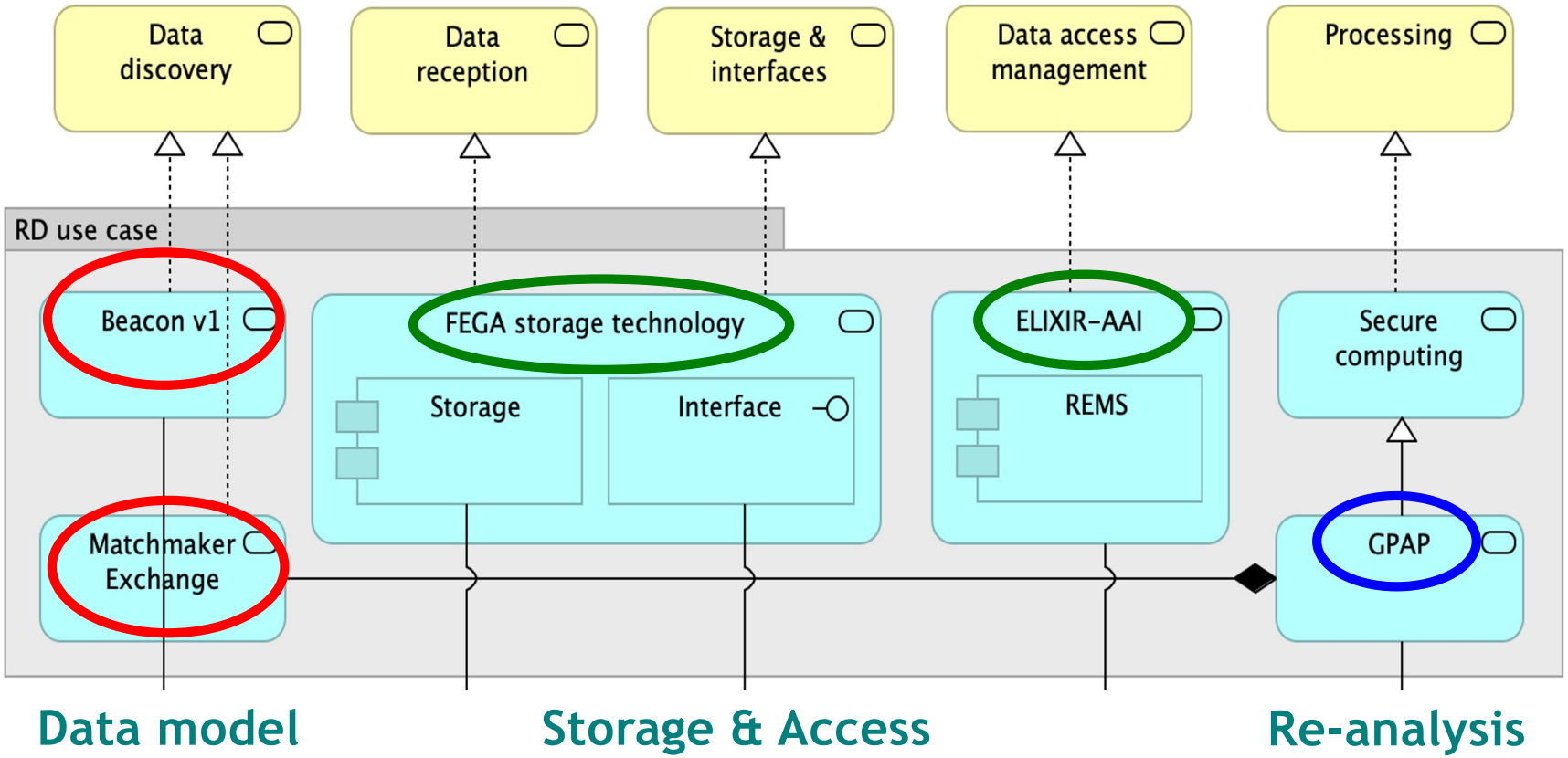
1+MG/B1MG working groups



ELSI & data governance framework

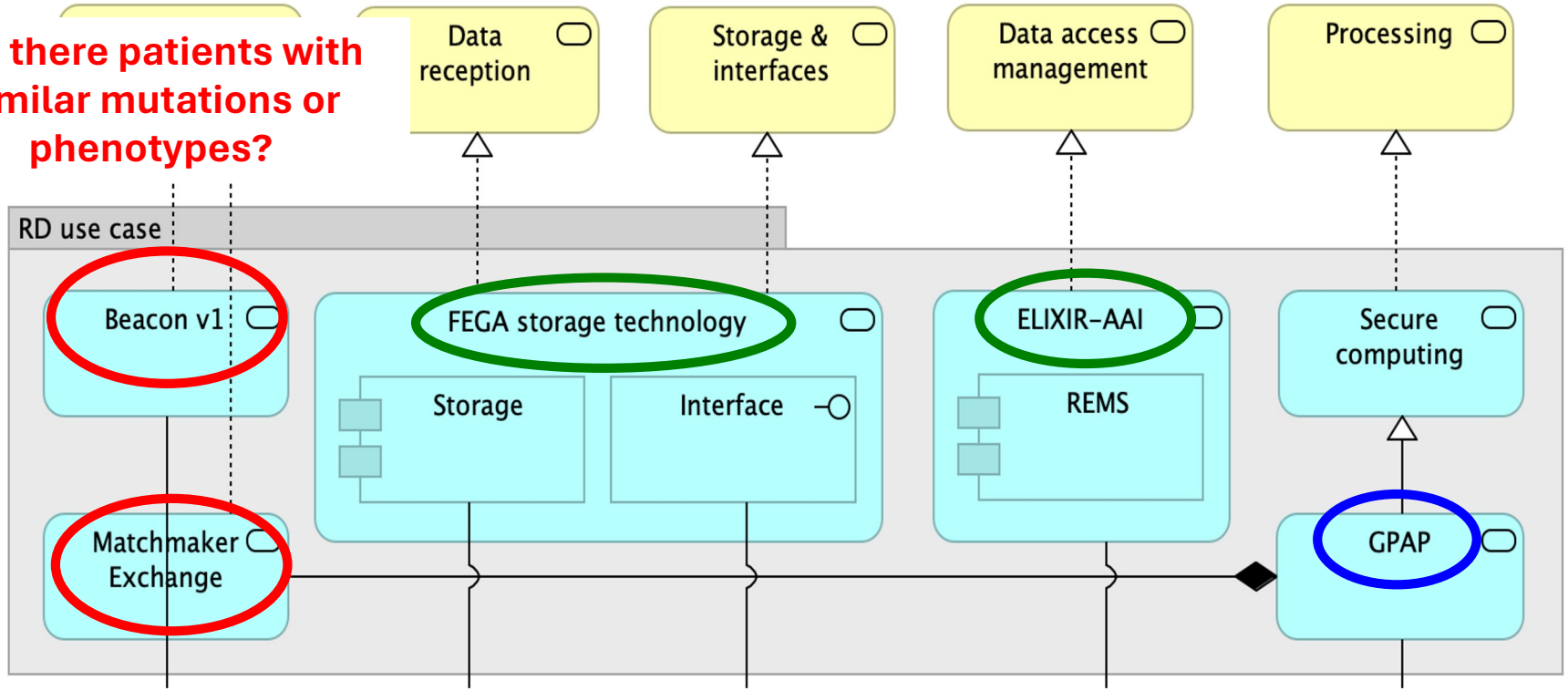


Proof of Concept(Rare Diseases)

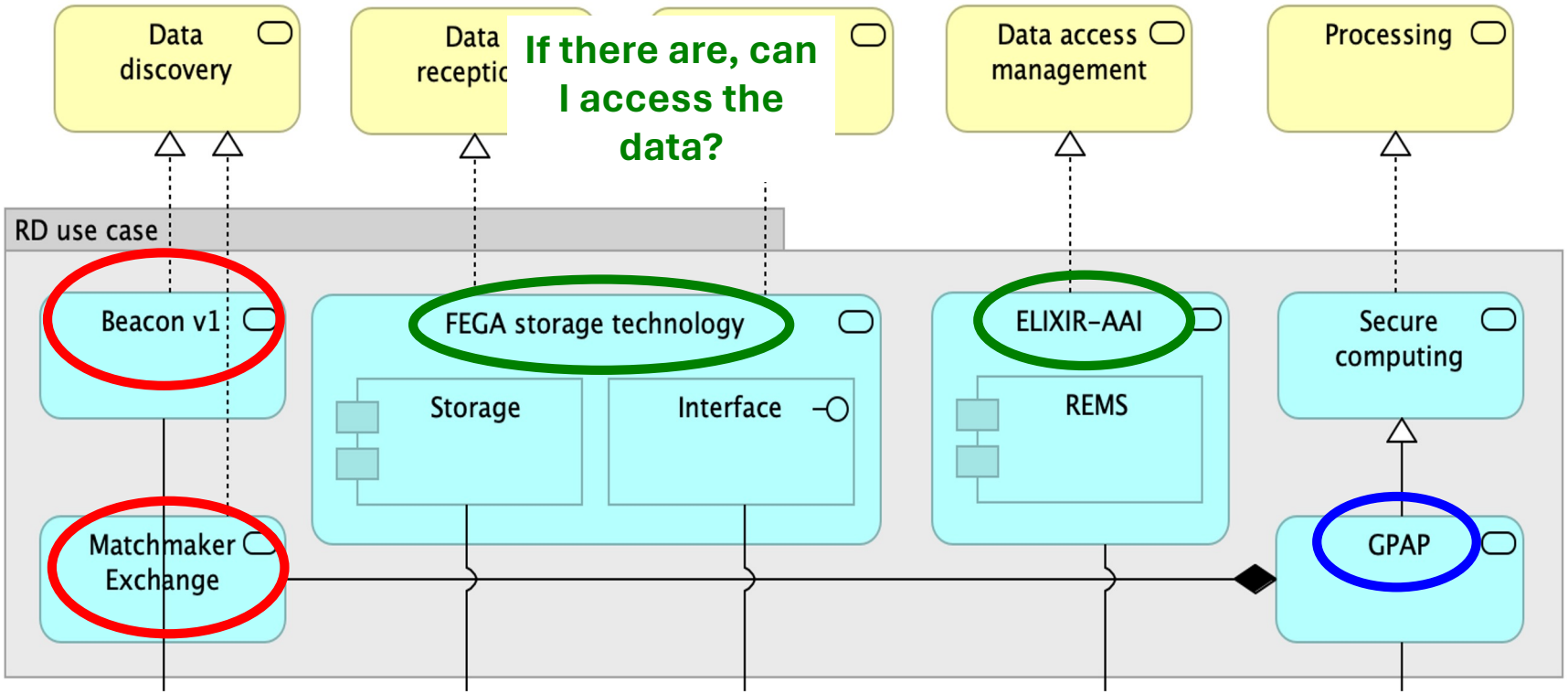


Proof of Concept(Rare Diseases)

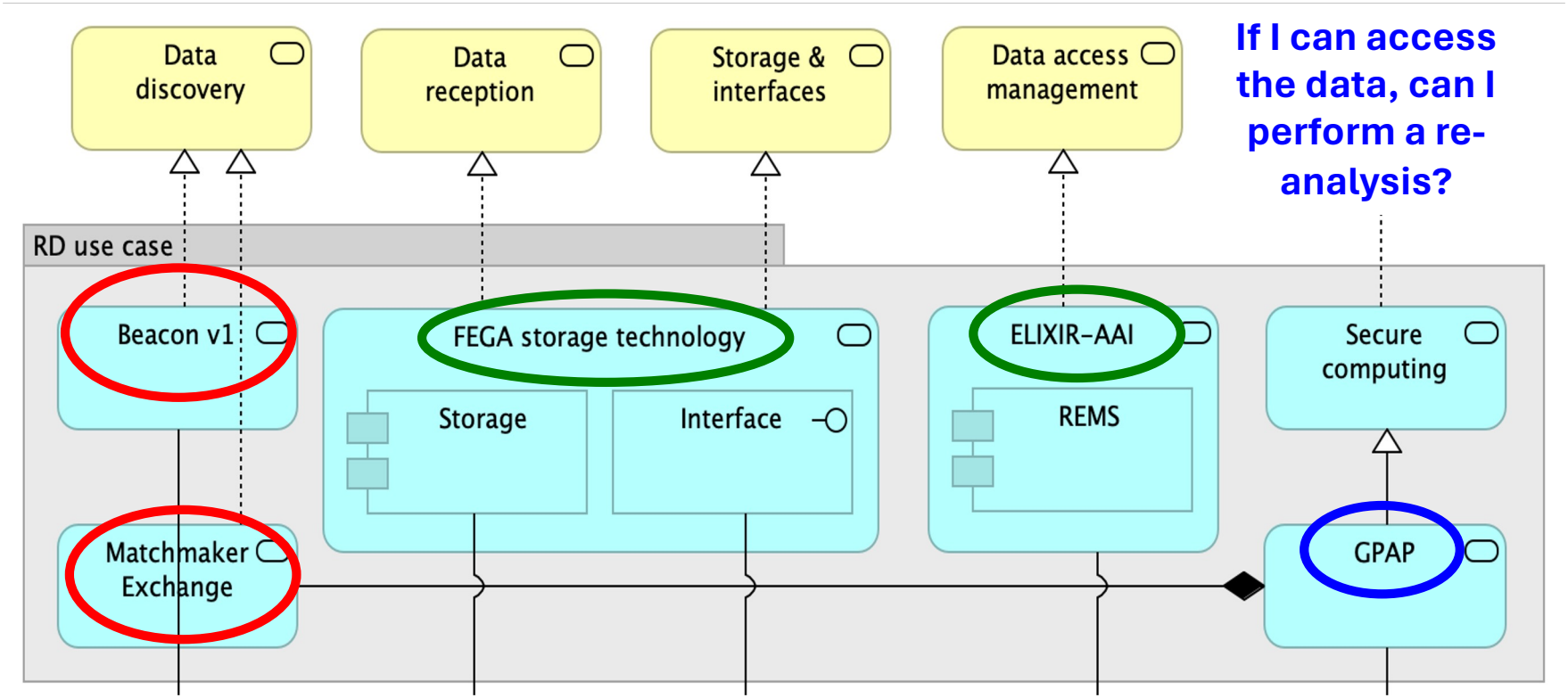
Are there patients with similar mutations or phenotypes?



Proof of Concept(Rare Diseases)



Proof of Concept(Rare Diseases)



1+MG Framework

Core 1+MG Framework

Technical framework

Sequencing guidelines

Sequence data generation and quality requirements for WGS/WES data to be labelled as 1+MG complaint

Data models, standards & ontologies

1+MG minimal data models for different use cases and recommendations on ontologies and data standards

Technical Infrastructure

Stack of standards, open source references implementations, synthetic data and proof of concepts that can be used to establish a 1+MG node.

Implementation

Data governance and ELSI

Guidance and recommendations on how to address data governance and ELSI aspect to ensure data can be made available.

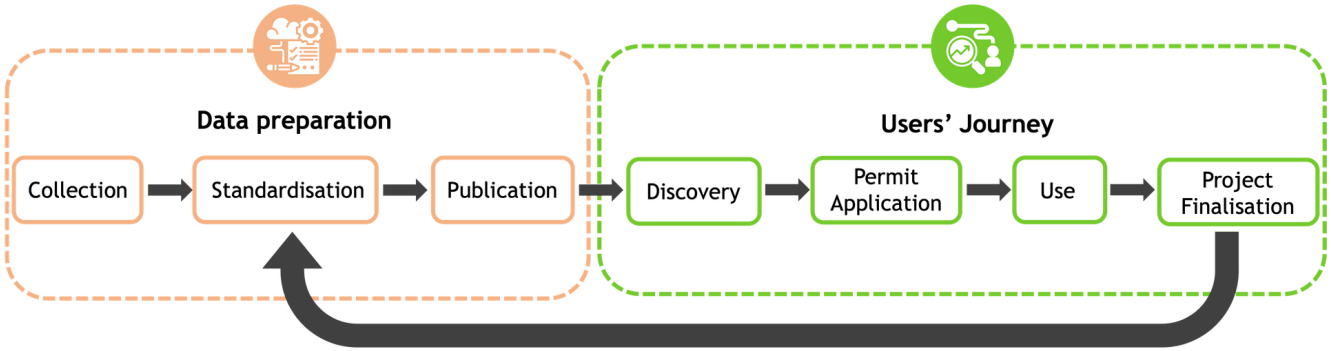
Genomics into healthcare

Assessment Maturity Level Model to guide healthcare systems on their journey to implement genomic medicine.

National implementation

Find pointers to country specific information resources and national research data management practices.

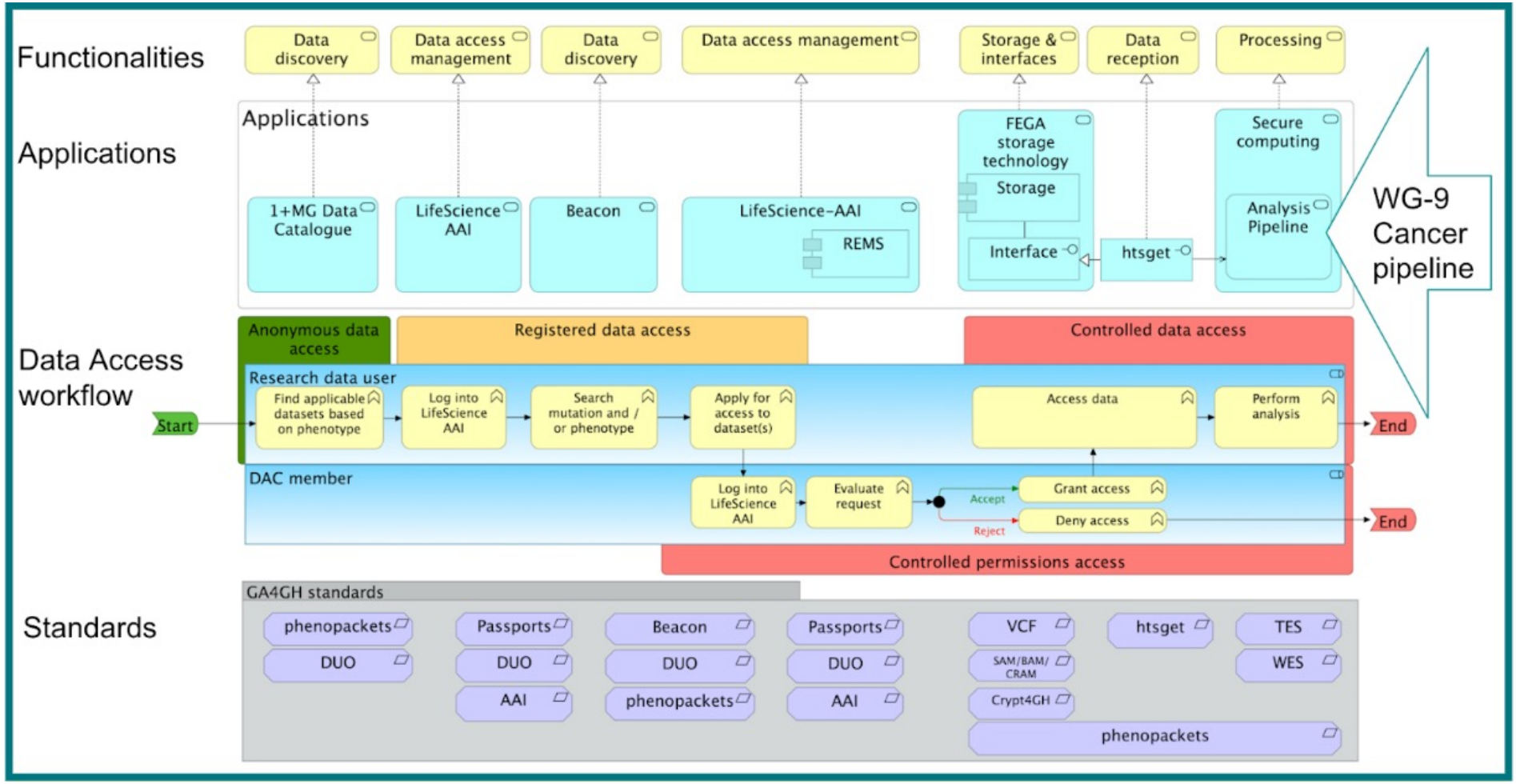
Alignment with EHDS



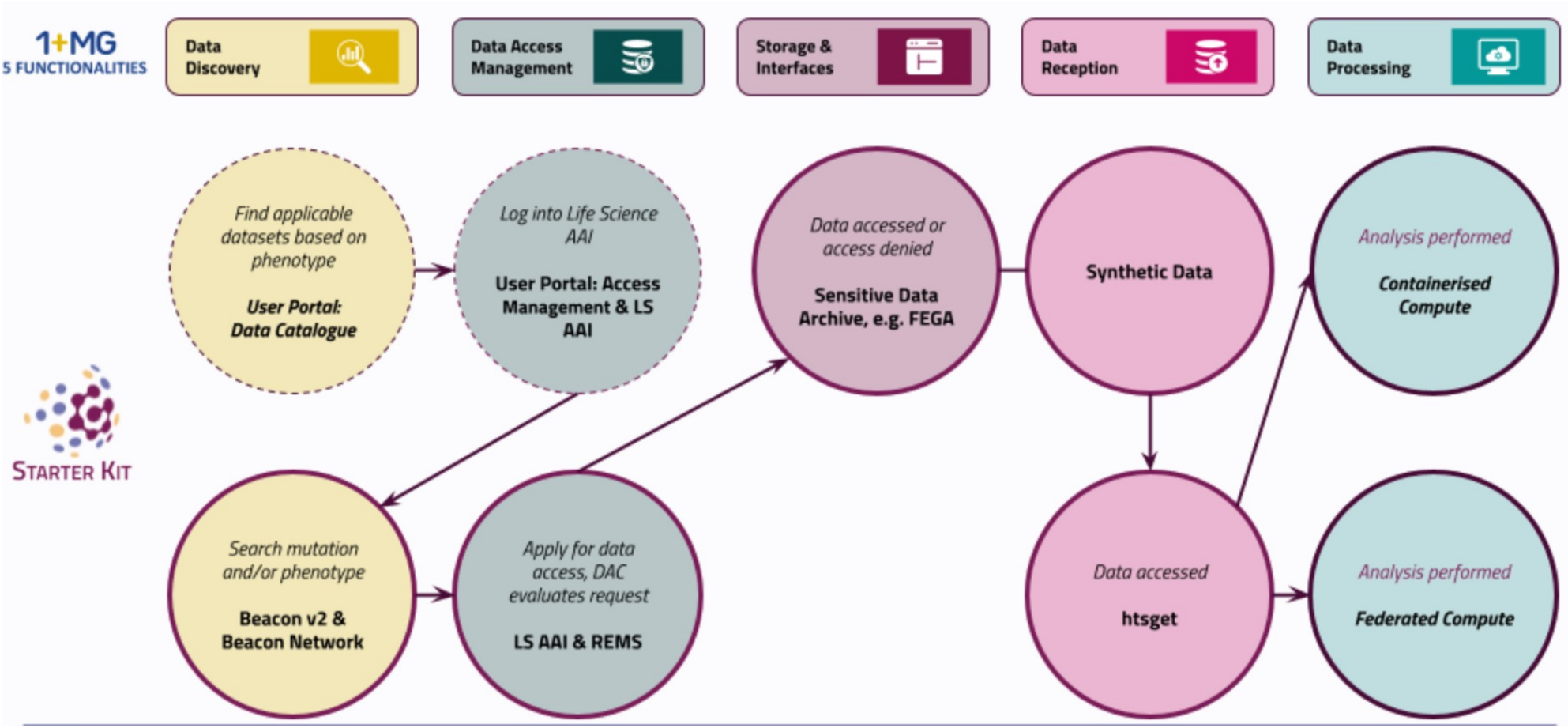
TEHDAS Proposed Data Lifecycle

EHDS HealthData@EU Pilot		1+MG Framework	
Data Preparation	Collection	Sequencing Guidelines	Governance & ELSI
	Standardization	Data Models & Ontologies	
		Data Reception	
	Publication	Data Storage & Management	
Users' Journey	Discovery	Data Discovery	
	Permit Application	Data Access Management	
	Use	Data Processing	
	Project Finalization		

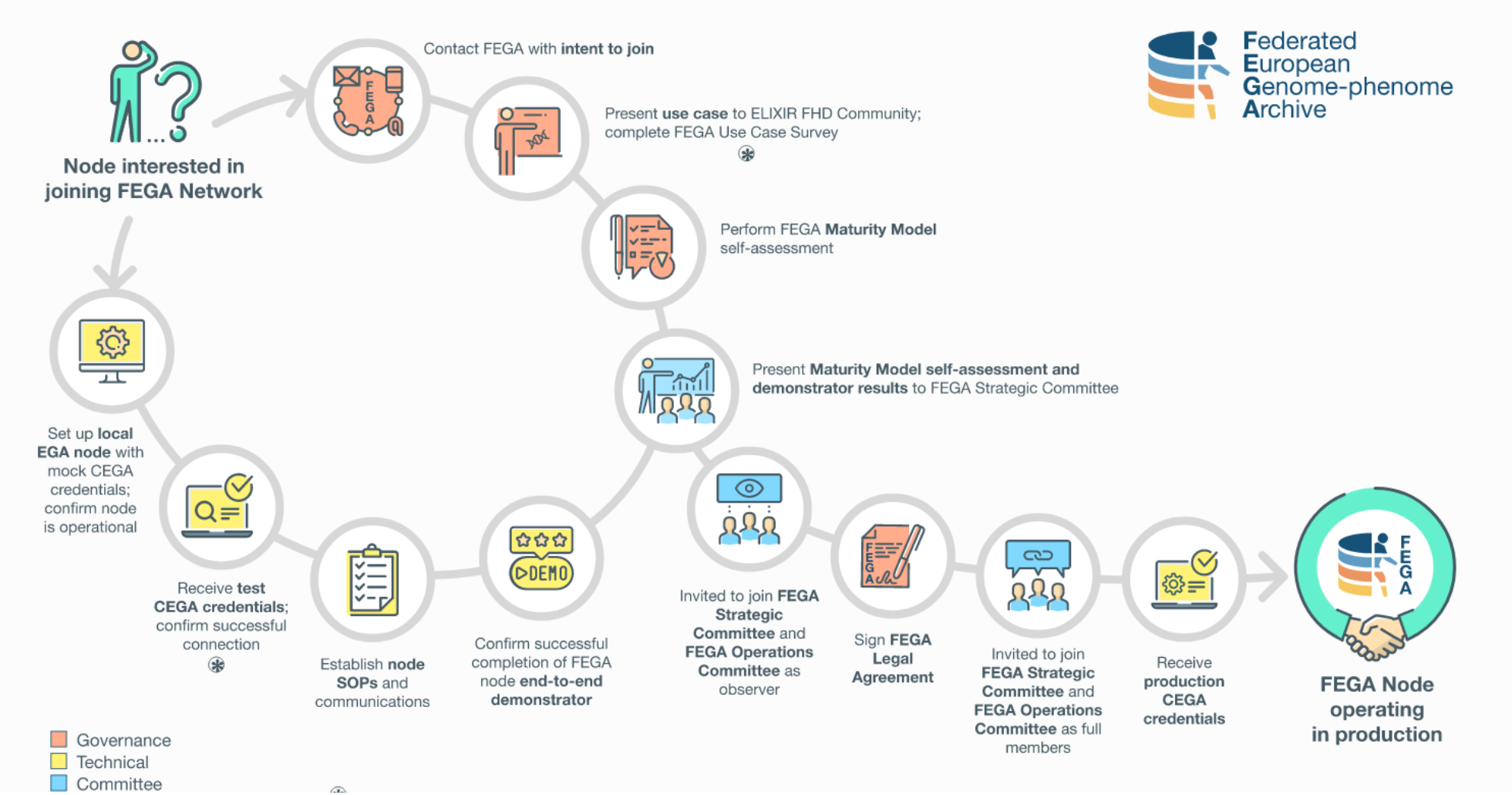
GDI evolution



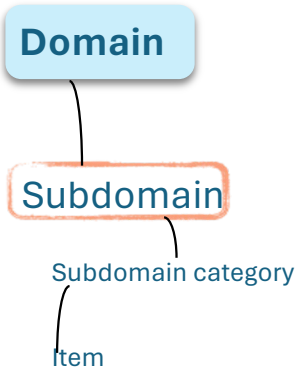
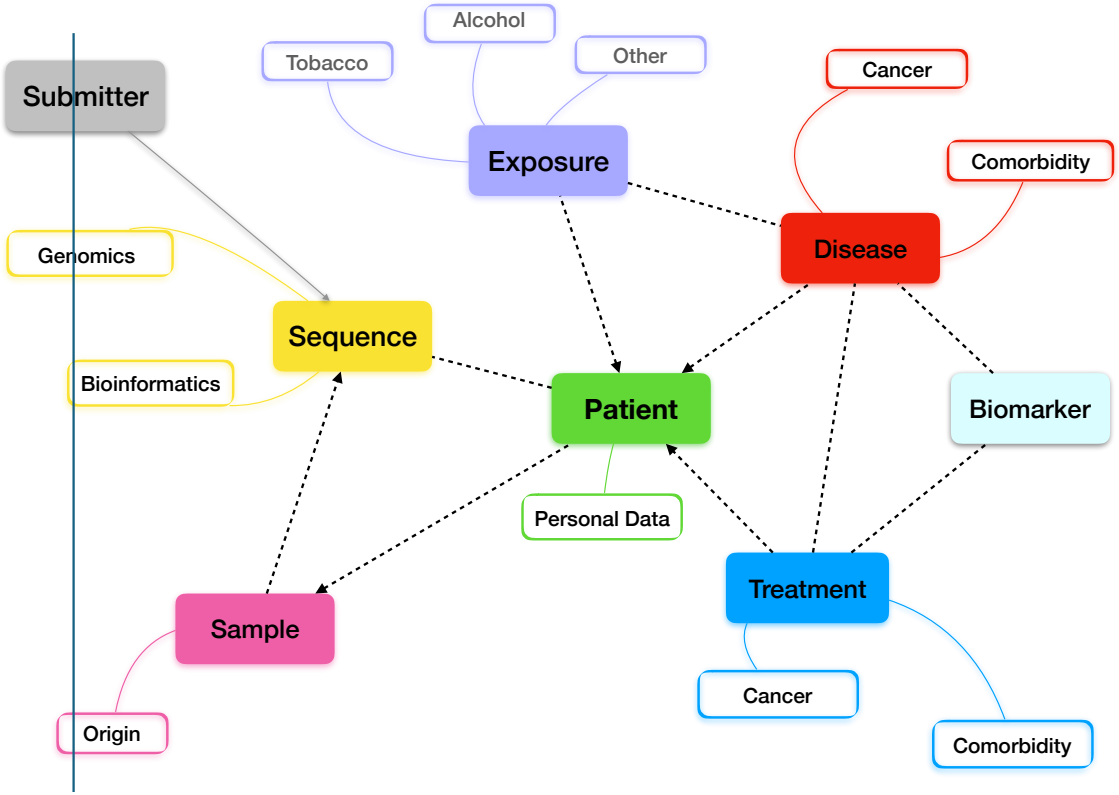
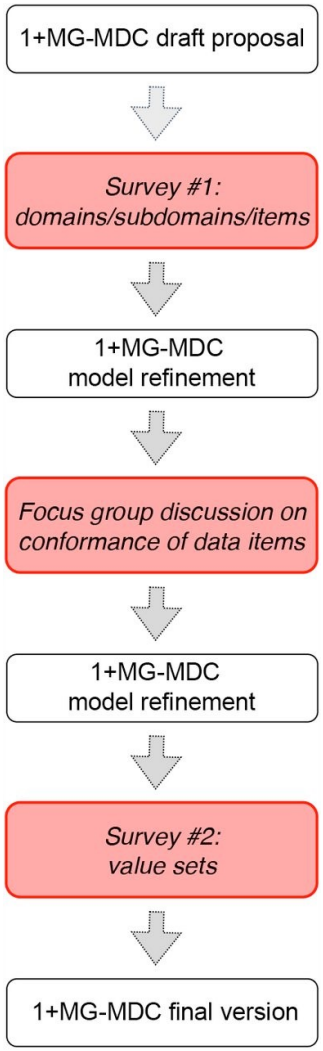
GDI Starter Kit



Federated EGA



Minimal Dataset for Cancer



- Mandatory
- Recommended
- Optional