EDITH's experience on the Virtual Human software Infrastructure

Sabato Mellone
University of Bologna

Amaryllis Raouzaiou
ATHENA RC

EDITH

EDITH is a coordination and support action funded by the Digital Europe program of the European Commission under grant agreement n° 101083771

To realise VHTs in Healthcare, work is needed in the following areas:

- **Technology**
  - Individual resources: data, models, algorithms
  - Integration of resources
  - APIs
  - Infrastructure
- **ELSI**
  - Access, privacy
  - Ethics, code of conduct
  - Legal & policy aspects
  - Regulatory considerations
- **Users**
  - Different profiles
  - Access & workflows
  - Interaction with other platforms & repositories
- **Sustainability**
  - Clinical uptake
  - Large companies
  - Start-ups
  - Marketplace
  - business modelling
  - ERIC, EDIH

VISION in EDITH

- A proof-of-concept operational repository of VHT resources
- A cloud-based federated infrastructure
- A catalogue of available VHT resources
- A tool to find, share, develop, curate, test/validate, connect together, compare, and support the regulatory acceptance/compliance of any VHT resource
- A tool to verify/ensure the compliance with established interoperability standards/profiles
- A tool to verify/ensure the compliance with EU HPC resources

STRATEGY in EDITH

- We defined technical requirements, legal requirements, desired integration/interoperability levels, and inclusion criteria.
- Inclusion criteria include accessibility, trustworthiness, legal/liability aspects, ethical aspects, compliance with FAIR data management principles, etc.

STRATEGY in EDITH: Trinity of Software

Catalogue

- A place to share and discover research objects (publications, data, models,…)
- Unique (global) identifiers & versioning
- Metadata (manual, automated) to facilitate discovery
- Actual object may be available (stored) elsewhere
- Catalogue and harvesting services (federation, distribution)

Repository

- Safely store and retrieve digital research objects (pubs, data, models,…)
- Files (versioning, metadata)
- Unique (global) identifiers & versioning
- Access policies (and sensitive data)
- Long-term preservation (storage classes)
- File & Sync (personal space)

Platform

- Analyze, simulate, visualize, process, manage, interact, …
- Software services (web apps, APIs, Jupyter notebooks, workflow engines, VDIs)
- Compute (HPC/HTC), storage and networking
- Collaboration and tiers
- Generic-purpose and domain-specific

Roles and responsibilities

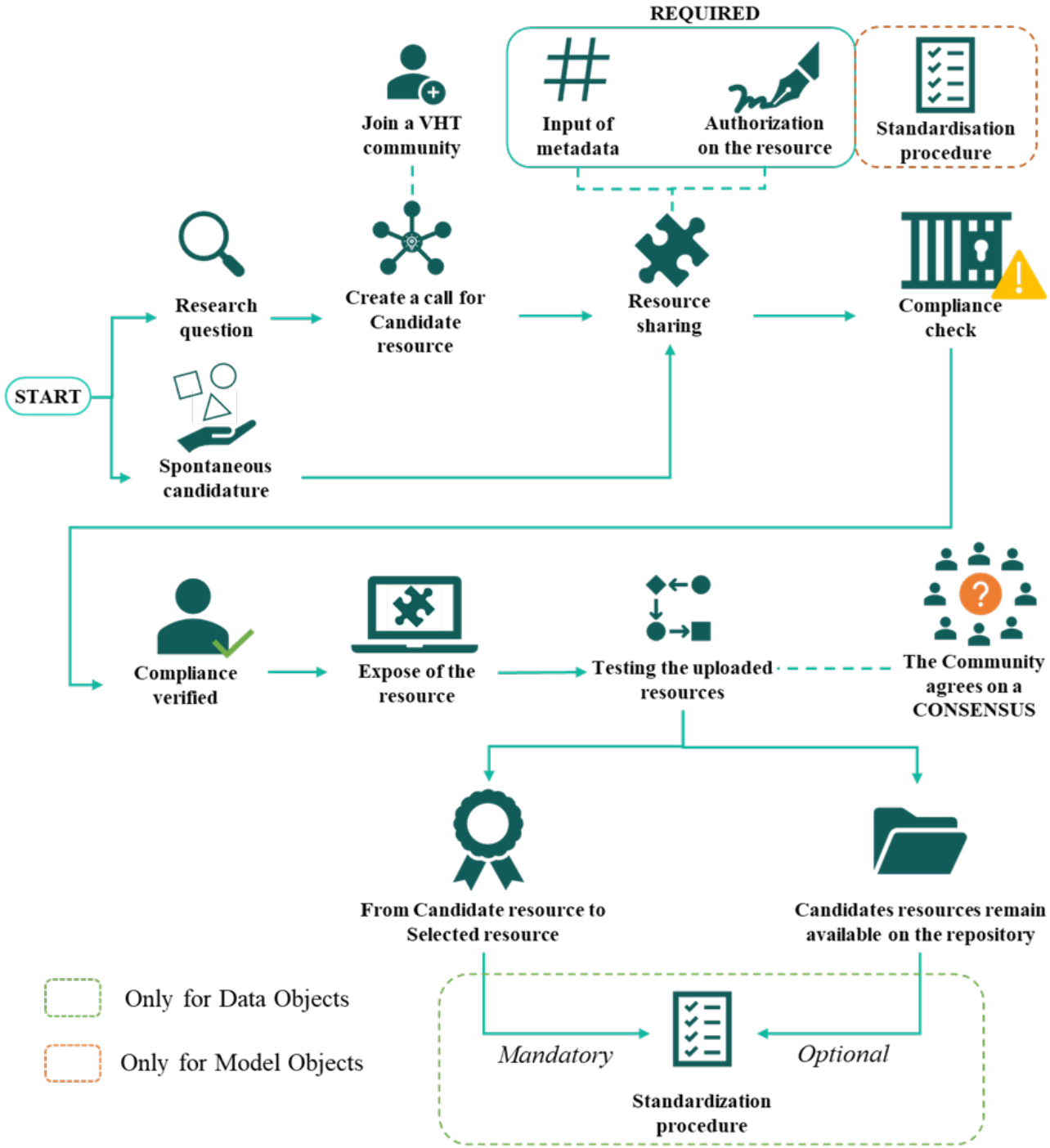
- In the VHT infrastructure, a **user profile** is defined as a collection of settings and information associated with a specific user.
- **User role** is a well-defined collection of capabilities within the system allowing the user to achieve the role's intended purposes.
- **Being a user of the infrastructure means becoming part of the community**

Roles and responsibilities

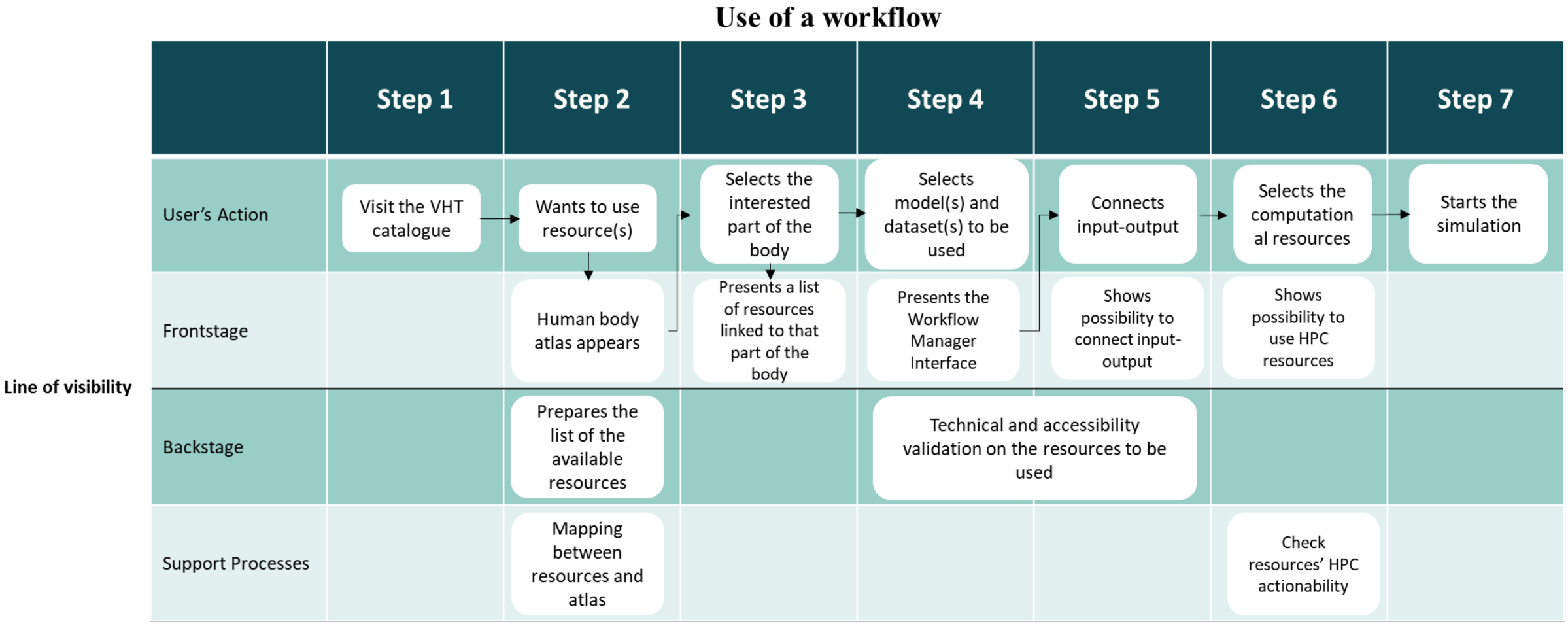
- Thinking at a high level, we can simply distinguish between providers, users, and “administrators”.
- **Providers**: must be authorised to make that resource available and must specify the terms of use
- **Users**: must be aware of the terms of use and comply with them
- **Administrators:** avoid making available unsuitable digital objects and/or with malicious intent. Ensure that Providers and Users can operate in agreement with their mutual obligations.

Decision-making processes

- The process of making a resource available within the VHT infrastructure, can start for at least two reasons
  a. A potential user browses the catalogue in search of a specific object and there is nothing like the one he/she wants;
  b. A provider wants to share their data/model object



Service blueprints



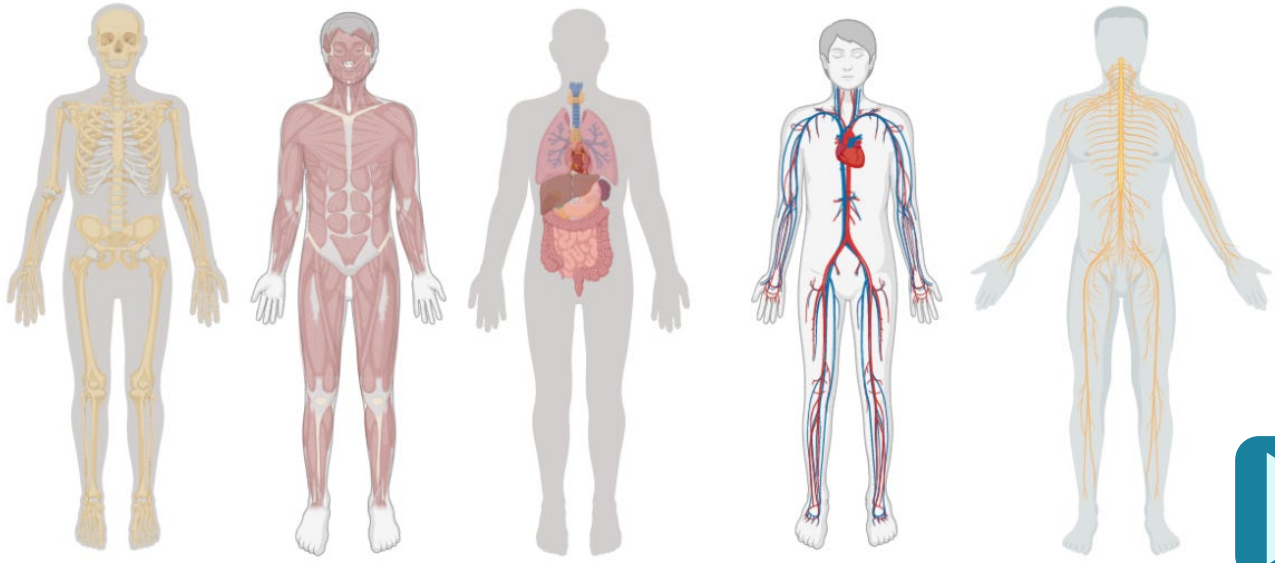
- Data Object Type
  - Semantics
  - Syntax
  - Accessibility
- Data Object Pose

Semantics

Select the scale of the dataset

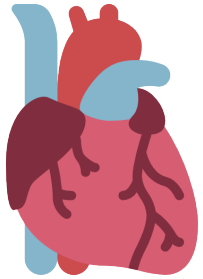
Molecular | Cell | Tissue | Organ | System | Body

Select to which part of the body the dataset is referred to



Use a resource

Select the part of the body to explore



Heart

Model #1
[Abstract]
INPUT: …
OUTPUT: …

Ontology: …
Accessibility: …
Body: …
Time (Age): …

Credibility: …
Clustering: …

Model #2
[Abstract]
INPUT: …
OUTPUT: …

Ontology: …
Accessibility: …
Body: …
Time (Age): …

Credibility: …
Clustering: …

Model #3
[Abstract]
INPUT: …
OUTPUT: …

Ontology: …
Accessibility: …
Body: …
Time (Age): …

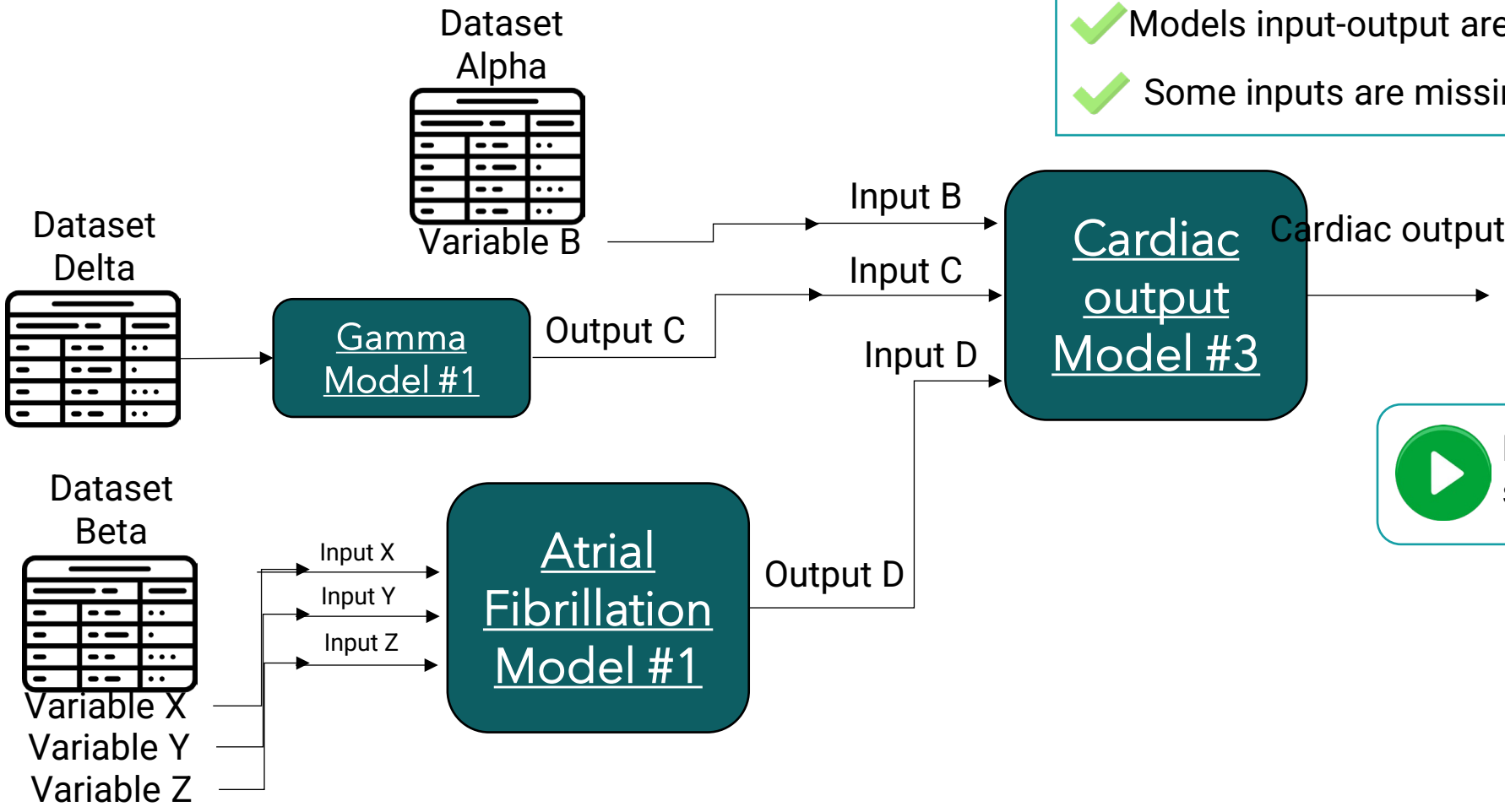
Credibility: …
Clustering: …

Cardiovascular system

Use a resource

Build the workflow

- All resources are actionable
- All resources are HPC-actionable
- Models input-output are compatible
- Some inputs are missing



Compliance and legal considerations

Legal considerations around the governance of the VHT infrastructure involve mainly three aspects:

- Privacy and Data Protection
- Accessibility of the resources (data/model sharing vs federation of resources)
- Licenses and terms of use

Compliance and legal considerations

Licenses and terms of use

Providers will indeed be asked to specify licenses and terms of use for their resources.

Users will have access to the resources provided by the third parties only when accepting the terms and conditions defined by the provider.

It is relevant to consider that when providers make visible and possibly available algorithms and software, Dir. 2009/24/EC on the legal protection of computer programs is applicable.

Developers

- early establishment of metadata schema
- semantic annotation to ensure interoperability (controlled vocabularies)
- model provenance tracking
- standardised formats (**non-negotiable**) for exchange and integration across platforms
- modular and composable model architecture
- continuous interoperability testing
- feedback loops between developers and contributors/stakeholders

Contributors

- modular model design — single-organ models adhere to standard interfaces to be composed into larger systems
- careful design of interfaces — clearly define model inputs, outputs, and boundary conditions, document units, data types, and assumptions
- ensure portability and reproducibility — information on dependencies, required libraries, and execution environments
- temporal/scale management — specify the model's temporal resolution, spatial granularity, and coupling requirements
- performance optimization — ensure models are computationally efficient and capable of running on HPC or cloud infrastructure
- standardised formats — to improve discoverability and interoperability
- controlled vocabularies — to maximize compatibility and reuse
- ethical, legal, and data protection requirements — ensure that data and models compl with GDPR and ethical standard

Stakeholders

Healthcare provider / clinician

- Interdisciplinary Collaboration
- Integration with Health IT Systems
- User-Friendly Interfaces
- Data Security and Privacy
- Training and Support

Patient

- Privacy and Security
- Easy registration and access
- Easy data uploading
- Training and Support

Researcher

- Sharing Models and Data
- Standardizing Research
- Version Control Systems
- Collaboration Platforms
- Modeling and Simulation