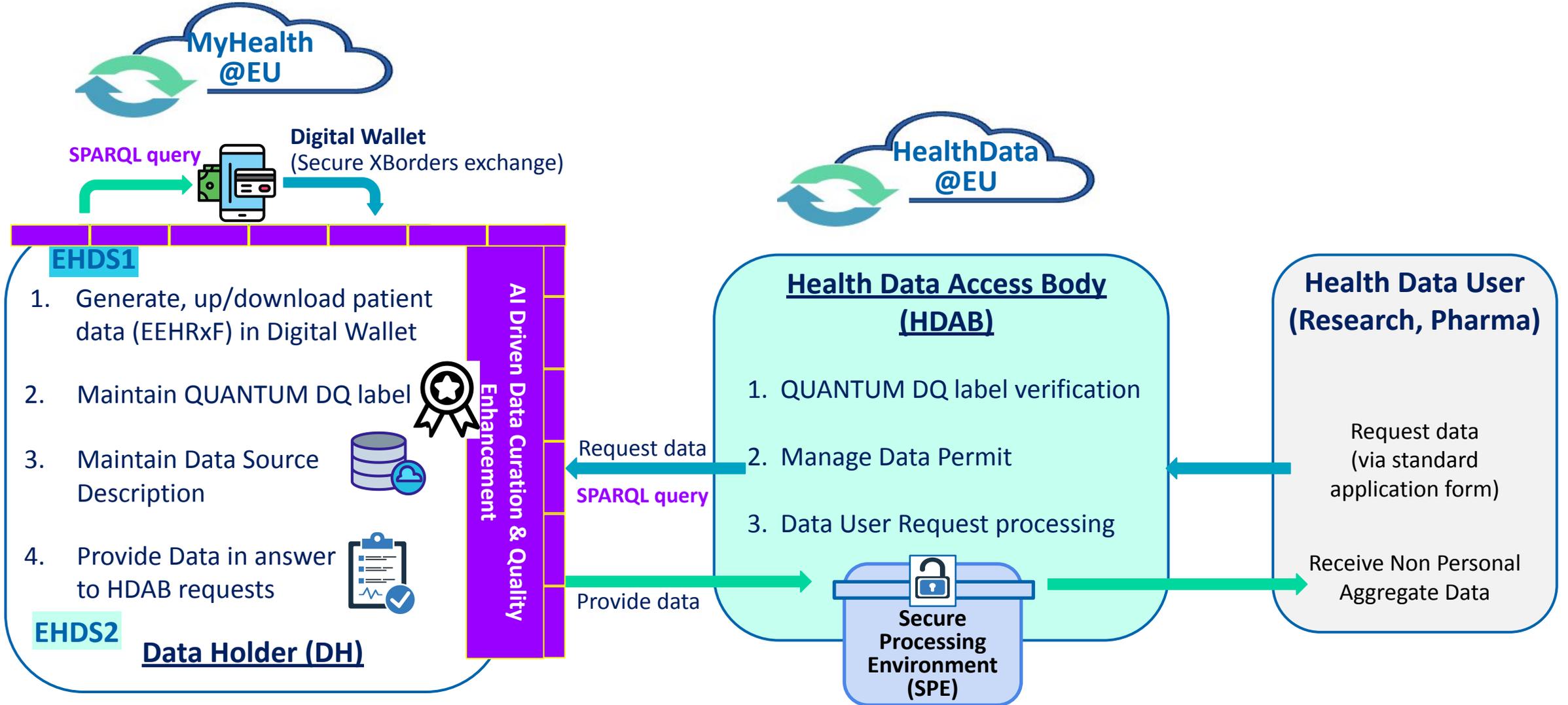# EHDS Data Flows with a 'Digital Twin' of patient health records

- EHDS data flows with a Digital Twin
- How to generate a Digital Twin with AIDAVA
- Benefits of Digital Twin
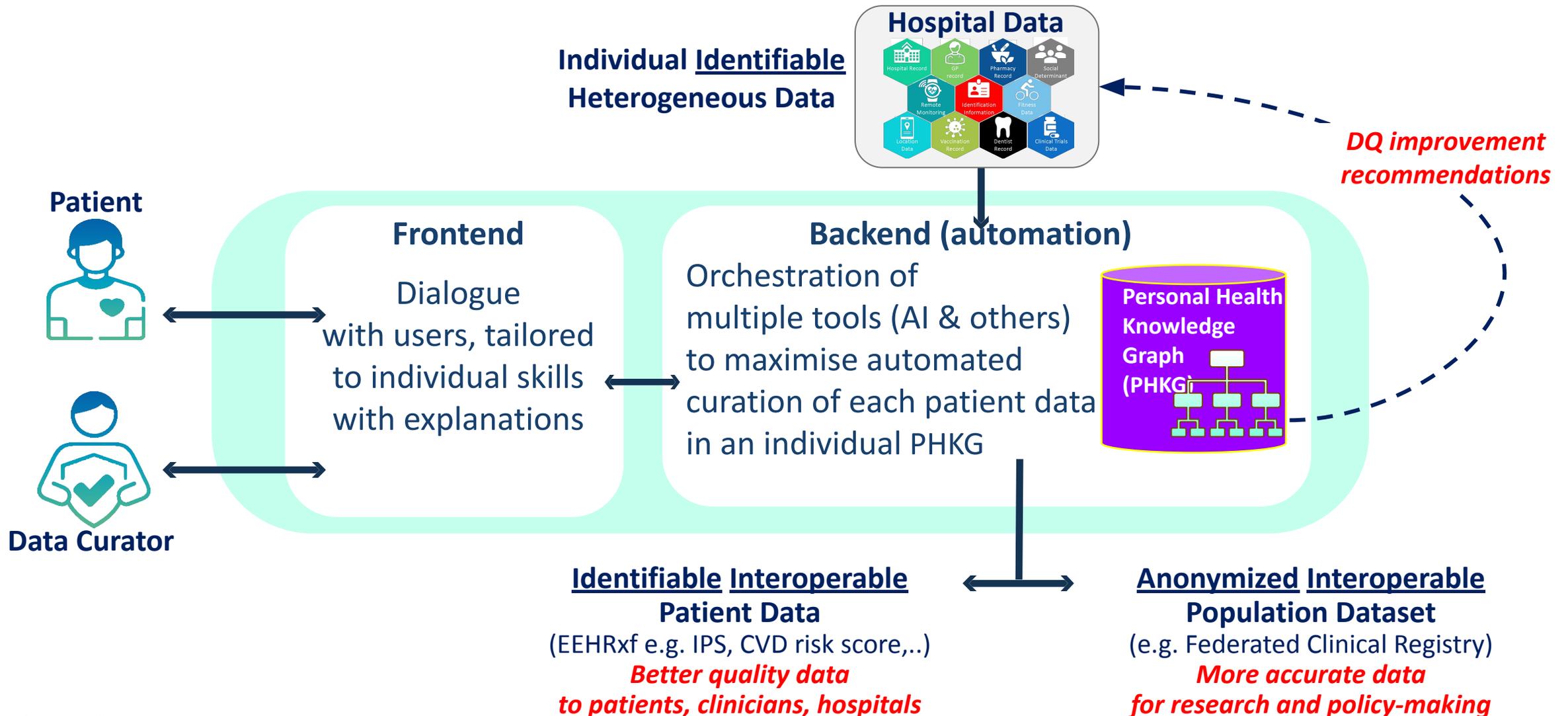- Critical Success Factors to AIDAVA Scalability

Deep dive in EHDS Data Flows

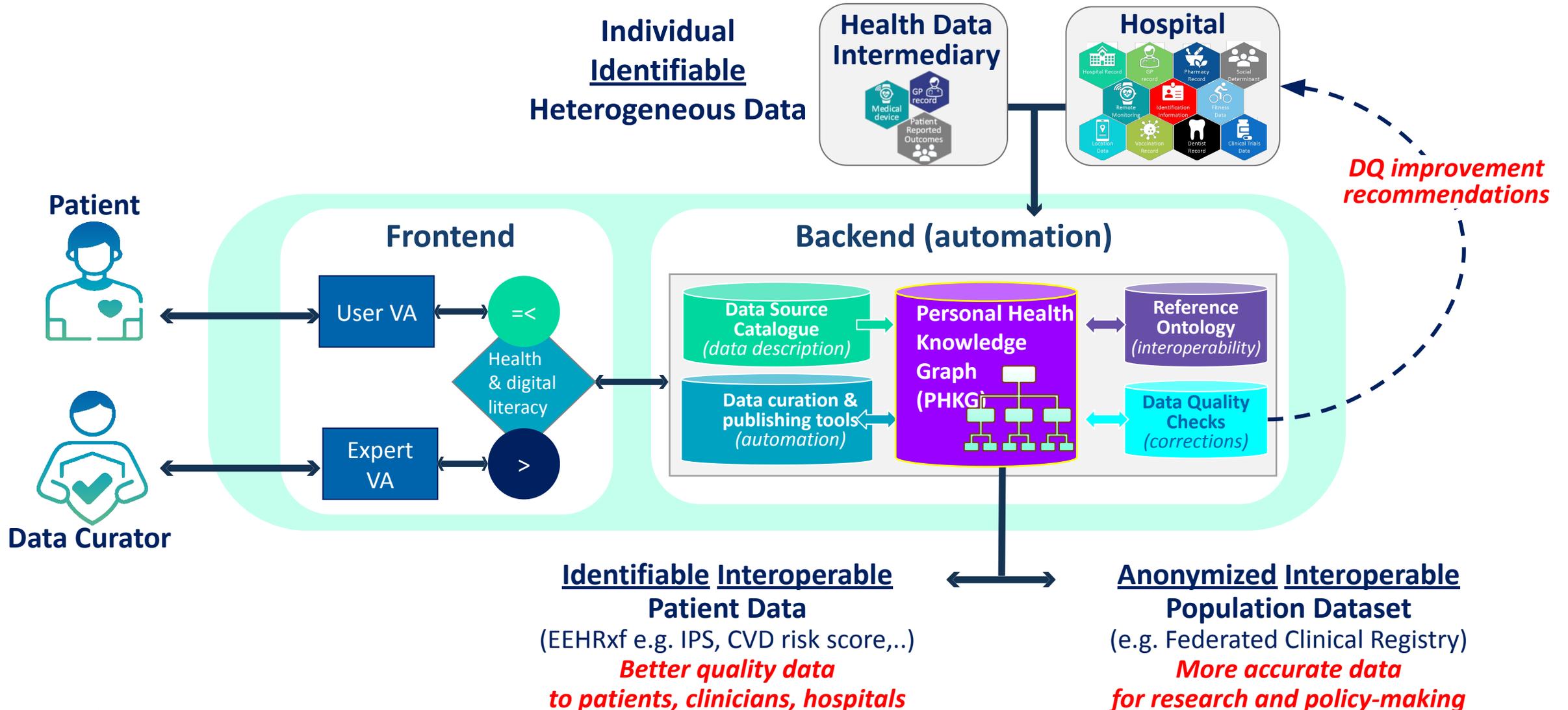## Adding AI maintained high-quality "digital twin" of patient records

MyHealth @EU

SPARQL query

Digital Wallet (Secure XBorders exchange)

AI Driven Data Curation & Quality Enhancement

EHDS1

Data Holder (DH)

1. Generate, up/download patient data (EEHRxF) in Digital Wallet
2. Maintain QUANTUM DQ label
3. Maintain Data Source Description
4. Provide Data in answer to HDAB requests

EHDS2

HealthData @EU

Health Data Access Body (HDAB)

1. QUANTUM DQ label verification
2. Manage Data Permit
3. Data User Request processing

Request data

SPARQL query

Provide data

Secure Processing Environment (SPE)

Health Data User (Research, Pharma)

Request data (via standard application form)

Receive Non Personal Aggregate Data

How to deliver a "digital twin" (PHKG)
Readiness for production use of AIDAVA

- AIDAVA prototype is a TRL 5-TRL 6 prototype
  - based on **21 AI and non AI tools** orchestrated through workflow to support automation in mapping, curation and data quality enhancement
  - available in **3 languages** (Estonian, German, Dutch)

- First generation (G1 - June 2024) tested with 78 patients in 4 hospitals
  - **documentation of available data** within an organisation is hard (knowledge & skills)
  - **automation in curation** is possible
    - dramatically decrease workload of current curation
    - increase interoperability of data (no risk of difference due to human mapping)
  - **major enhancements** required
    - AI tools (NLP, medical coding, mapping) across all languages
    - Explainability (context, language) in case of issues

- G2 (December 2025) being tested with 90 patients in 4 hospitals
  - Serious improvements to AI tools (more is needed)

Benefits of high quality data "digital twin"
Quality = interoperable, correct/accurate, complete AND reusable

EHDS compliance at lower cost for Data Holders
- Smoother answer to queries: HDAB data requests ("permit") transformed in SPARQL query executable everywhere without additional transformation
- Easier generation of EEHRxF : implementation guide for each critical data category translated into SPARQL query executable everywhere
- Limited need for new skills/staff: data scientists to support AI driven curation

Additional benefits for Data Holders: Data quality at INDIVIDUAL record level
- Primary use: Improved patient outcome and personalised (AI) care: individual data in AI model ready format (extracted from digital twin)
- Secondary use: Automatic maintenance of local (or federated) clinical registries and smoother participation to Clinical Trials for all hospitals (including rural ones)

## Additional benefits for Authorities: Faster & more reliable evidence
- Higher quality health systems and public health strategies,
- Faster and more impactful innovations adapted to all EU population

How to deliver "digital twin" (PHKG) at EU level
Critical Success Factors (CSF)

- AIDAVA demonstrated that (mainly) automated curation is possible by orchestrating multiple AI and non-AI tools
- AIDAVA-**like** systems could be deployed across EU with 2 CSF
  - Agreement & enforcement on **SULO** (Simplified Upper Level Ontology)
    - SULO supports alignment
      NOT harmonization across standards
      - by making their semantic explicit
      - and facilitate (AI driven) mapping
    - All "Digital Twins" should be compliant with SULO
      (no impact on local HIMS/EHR systems)
  - Speed of improvement of **AI tools** across EU languages
    (mainly NLP and Entity Linking/medical coding)
  - Availability of detailed **documentation on data sources**
    (extension of EHDS Data Description to each attribute)
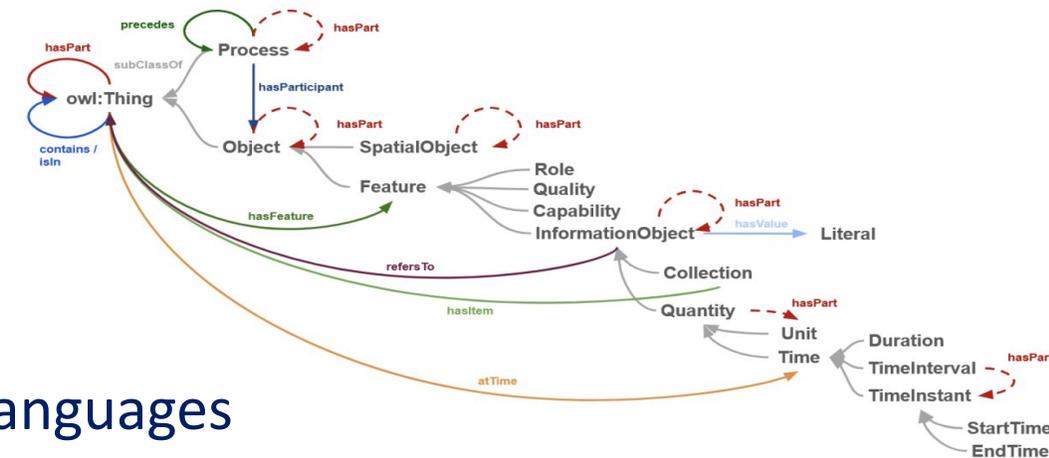
Comparative cost estimate of EHDS - 2 scenarios [ref]
10 hospitals, 5 years, per year: 1250 records, 1250 EEHRxF, 100 HDAB queries

	Steps	CURRENT SCENARIO (K€)					Curation of data sources - with AI (K€)			
		One of	Per unit	Yearly	Total Period		One of	Per unit	Yearly	Total Period
Data Holders/ EHR Vendors	Curation	€0	€0	€0	€0		€300	€0.10	€425.0	€2,125
	EEHRxf	€300	€0.10	€425	€2,125		€20	€0.05	€82.5	€413
	Data Description	€50	€0.25	€52	€258		€5	€0.05	€5.3	€27
	Answer to permit	€0	€5	€500	€2,500		€10	€0.05	€15.0	€75
	Total/Data Holder				€4,883					€2,639
	TOTAL ALL DH				**€48,825**					€26,390
HDAB	Curation	€0	€50	€5,000	€25,000		€0	€5	€500	€2,500
TOTAL					**€73,825**					**€28,890 (39%)**

34

Comparative cost estimate of EHDS - 2 scenarios [ref]
10 hospitals, 5 years, per year: 1250 records, 1250 EEHRxF, **200 HDAB queries**

	Steps	CURRENT SCENARIO (K€)				Curation of data sources - with AI (K€)			
		One of	Per unit	Yearly	Total Period	One of	Per unit	Yearly	Total Period
Data Holders/ EHR Vendors	Curation	€0	€0	€0	€0	€300	€0.10	€425.0	€2,125
	EEHRxf	€300	€0.10	€425	€2,125	€20	€0.05	€82.5	€413
	Data Description	€50	€0.25	€52	€258	€5	€0.05	€5.3	€27
	Answer to permit	€0	€5	€1,000	€5,000	€10	€0.05	€15.0	€100
	Total/Data Holder				€7,383				€2,664
	TOTAL ALL DH				**€73,825**				**€26,640**
HDAB	Curation	€0	€50	€10,000	€50,000	€0	€5	€1,000	€5,000
TOTAL					**€123,825**				**€31,640 (25%)**