InteropEHRate

EHR in people's hands across Europe

# SEMANTIC TOOLS AND METHODOLOGY FOR HEALTHCARE DATA INTEROPERABILITY

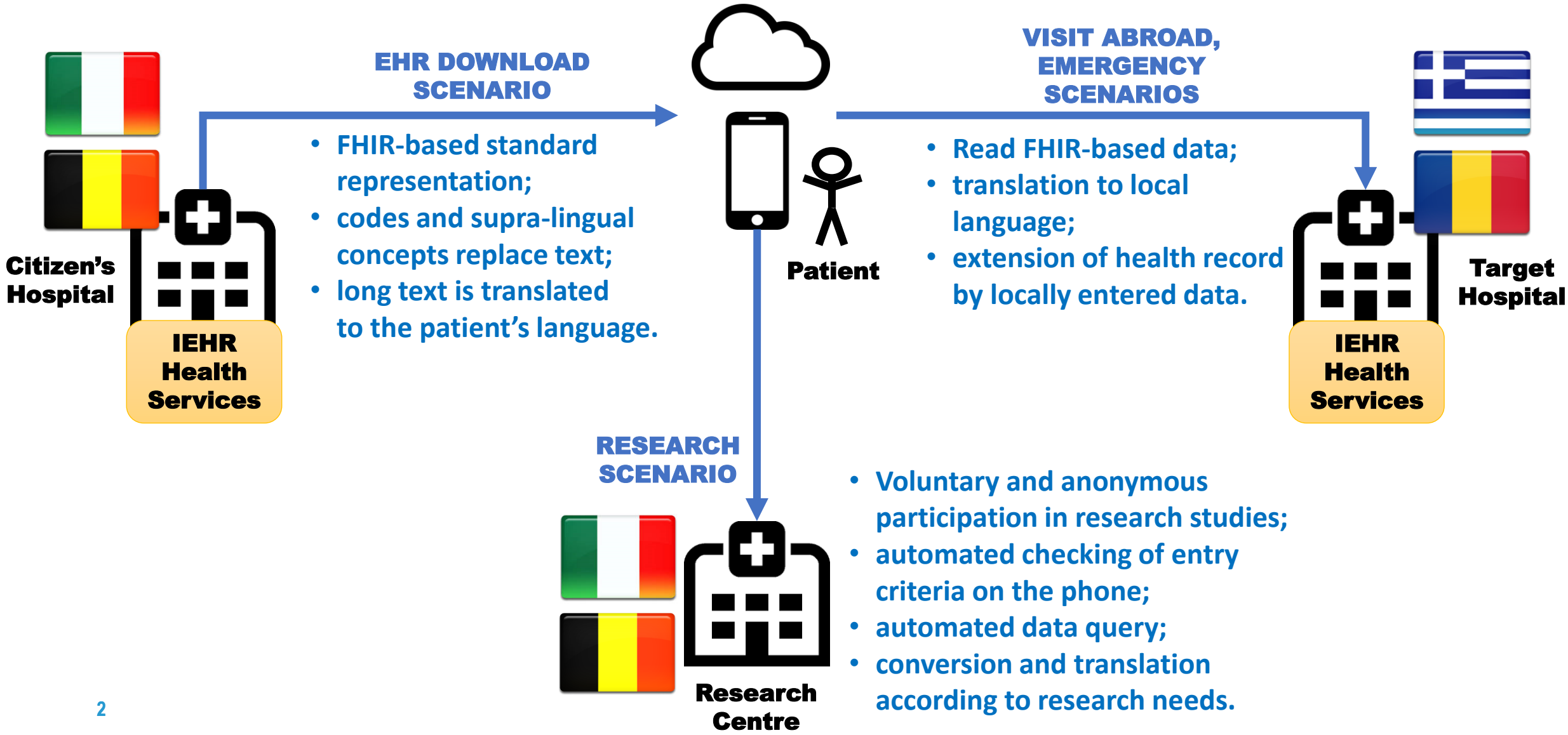
**SIMONE BOCCA, GÁBOR BELLA, YAMINI CHANDRASHEKAR**

FINAL CONFERENCE

28 SEPTEMBER 2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826106

# THE INTEROPEHRATE PROJECT



# EXISTING SOLUTIONS

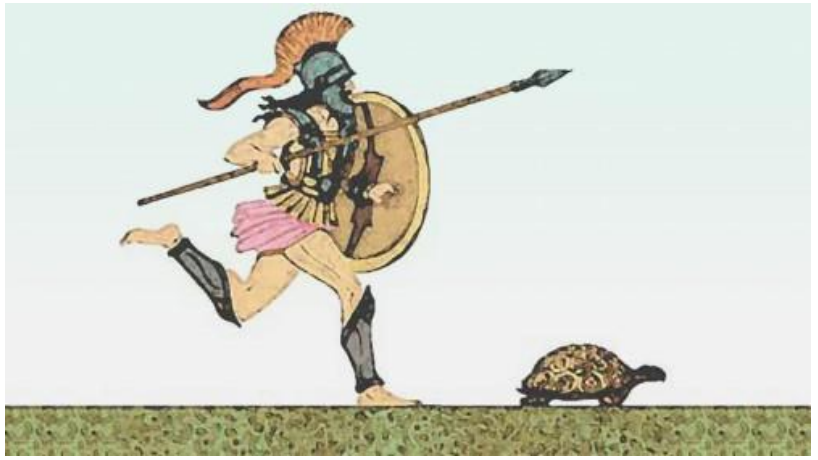
## National and international standards:

- monolingual and multilingual **terminologies** (SNOMED CT, UCUM, EDQM);
- international **coding systems** (ICD, LOINC, ATC);
- **information models** (HL7, CDA, FHIR, OMOP);
- there is progress on adoption on both local and governmental levels, but it remains limited.

## Supporting technology:

- **terminology servers**: increasingly used;
- **Extract-Transform-Load (ETL) tools**: graphical UI for file conversions, schema mappings, data transformations, sometimes code mappings: widely used;
- **information extraction tools**: never in care, somewhat in research, more often for accounting.

# CHALLENGES



## ADOPTING STANDARDS IS HARD

Healthcare standards are very complex

=> it is easier to build custom ad-hoc systems than full support for e.g. FHIR.

## HEALTHCARE IS EVOLVING

Standards, tools, and even our knowledge of healthcare is constantly changing

=> the system must be agile to support this evolution.

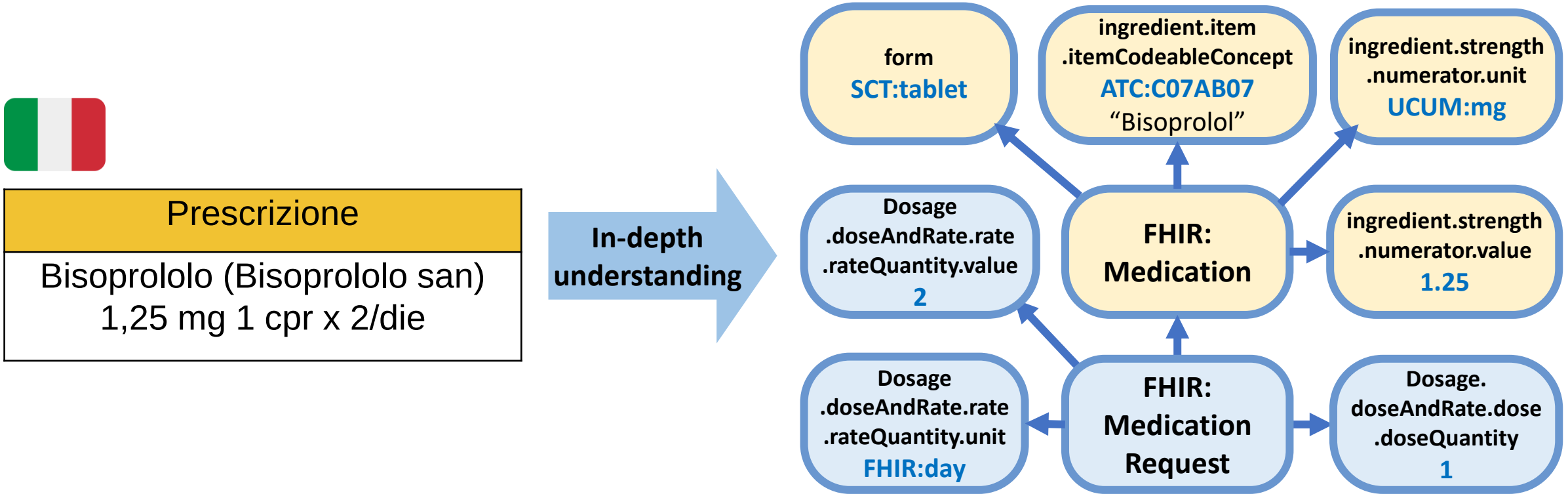
## AUTOMATION REDUCES PRECISION (AND VICE VERSA)

Care and research require ~100% precision and explainability in data processing.

But human input does not scale well over large quantities of data.

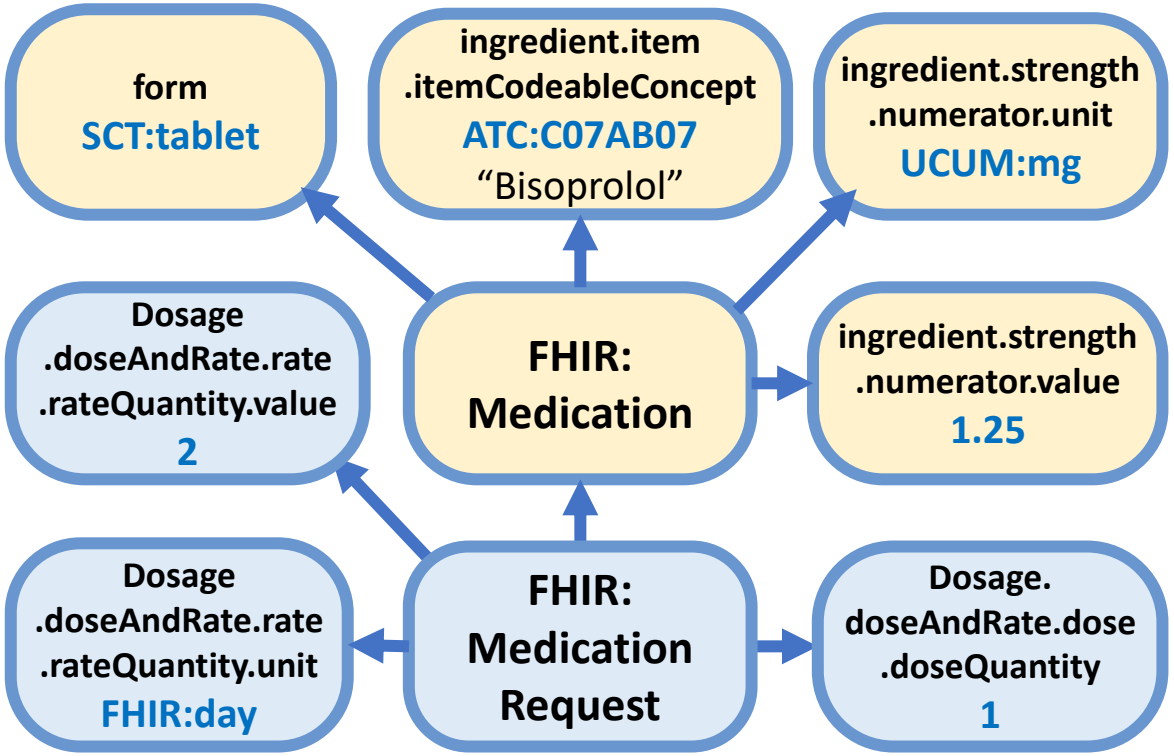
# DEEP SEMANTIC INTEGRATION

"Deep and semantic" means that every single data value (relevant to the task) is understood, made explicit, and represented in a formal, language-independent manner.



# DEEP SEMANTIC INTEGRATION

Starting from a correct and complete knowledge graph, **in-depth adaptation** (conversion, translation) can be automated in a robust way.

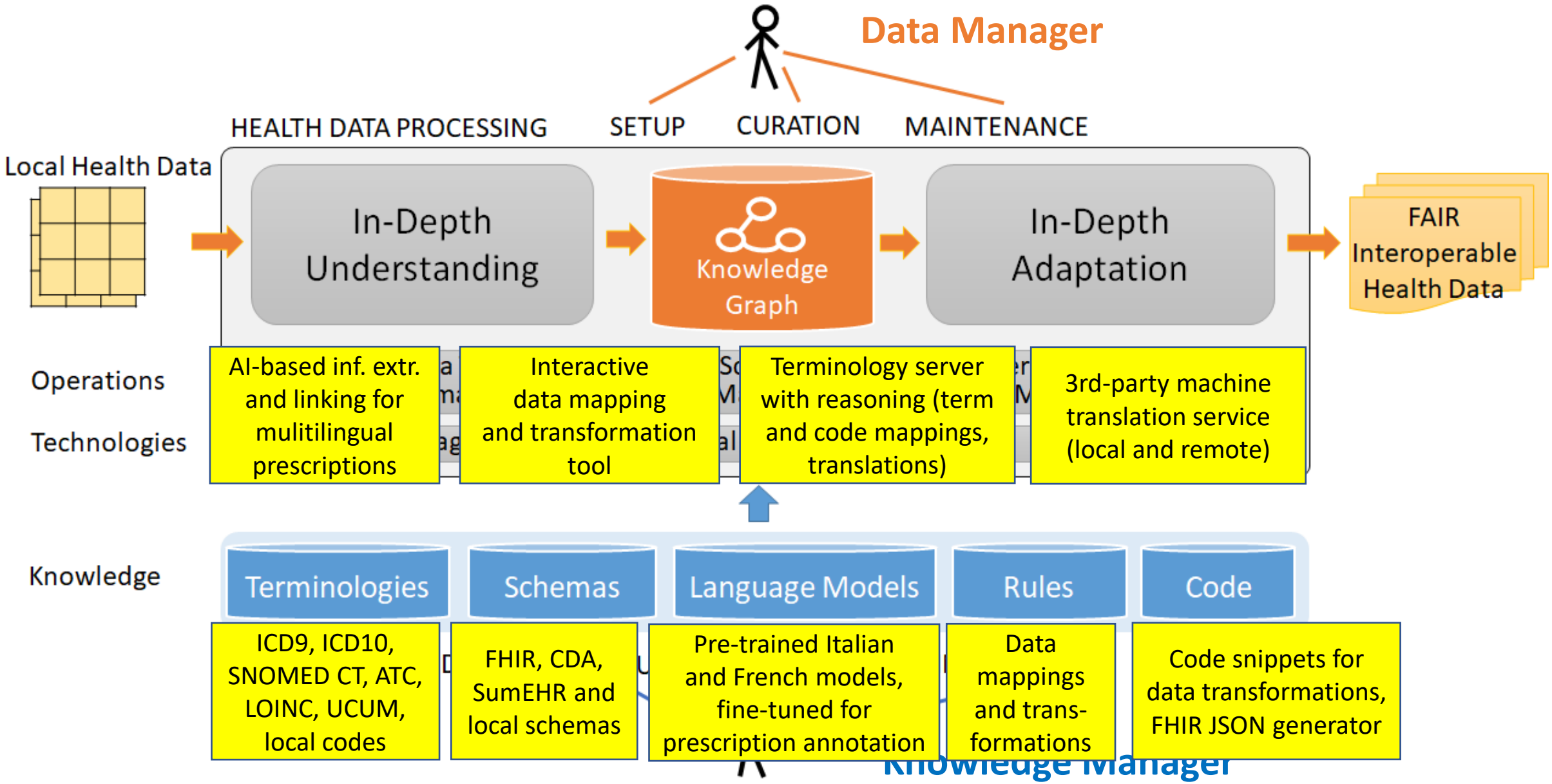



In-depth adaptation

| Principe actif | Produit | Dosage | Unité de mesure |
|---|---|---|---|
| Bisoprolol | Bisoprololo san | 1,25 | mg |

| Quantité | Forme pharmaceutique | Fréquence | Période |
|---|---|---|---|
| 1 | comprimé | 2 | jour |

# COMPONENTS AND HIGH-LEVEL METHODOLOGY



# A HUMAN-CENTRED, YET AUTOMATED METHODOLOGY

For correctness, **human supervision is essential**.

- **Knowledge manager**: setup, curation, and maintenance of **knowledge** (term bases, schemas, language models, rules, code snippets).
- **Data manager:** setup, curation, and maintenance of the **data mapping and transformation** process.

How to reconcile human supervision and scalability?

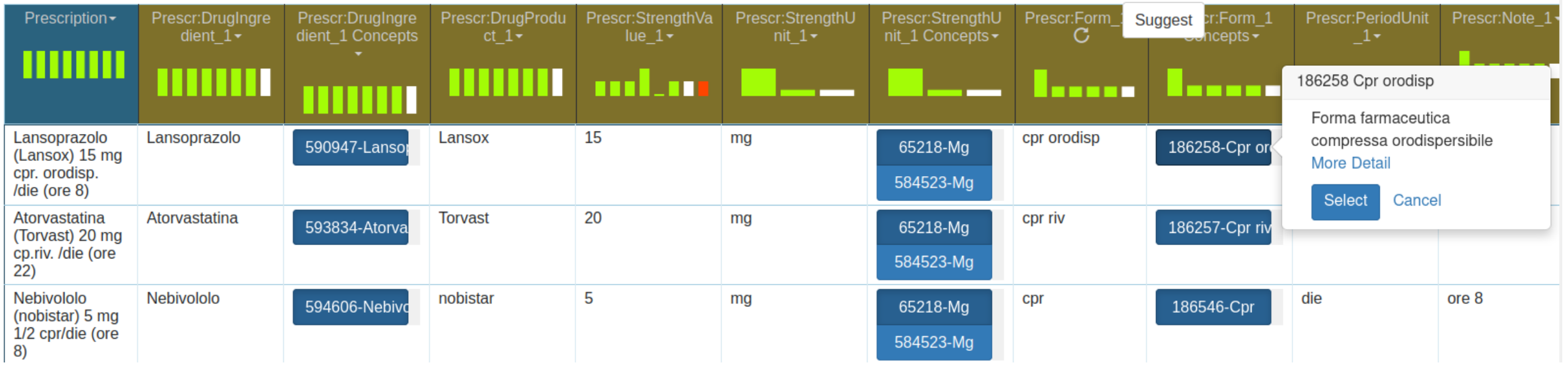
Through a **set of knowledge and data management tools**, integrated into a single GUI.

# INNOVATIVE TOOLS AND METHODOLOGY

The graphical ETL paradigm is extended by

**multilingual information extraction** and **knowledge graph building**.

The InteropEHRate semantic health tools have been primed twice by the EC Innovation Radar as outstanding innovation.



# IN-USE EXPERIENCE FROM INTEROPEHRATE







| | | | |
|---|---|---|---|
| **NATURAL-LANGUAGE TEXT** | Italian | French | Italian, French, Greek, Romanian |
| | Codification of NL terms, information extraction from prescriptions | | Official translations of code definitions, machine translation of longer text |
| **TERMS AND CODES** | HL7 service codes, ICD9-CM, LOINC, ATC | Local codes, ICD10-CM, ATC | ICD10-CM, LOINC, ATC, SNOMED CT, UCUM |
| **SCHEMAS** | CDA | SumEHR (IPS) + local schemas | FHIR directly used by HCP Application |
| **FHIR RESOURCES** | Patient, Practitioner, Encounter, Condition, CarePlan, Medication, MedicationStatement, Observation, DiagnosticReport, Media, AllergyIntolerance | | |

# CONCLUSIONS AND LESSONS LEARNT FROM PILOTS

- Over health data, **full interoperability is hard** and will always be.
- Standards and supporting tools are necessary but **not sufficient**.
- We try to reconcile precision and automation through **methodological innovation**.
- In the InteropEHRate pilots, **human effort was still crucial**:
  - **local data experts** (Fondazione Monasterio, CHU de Liège);
  - **methodology expert** (University of Trento);
  - **standardisation expert** (Fraunhofer Institute).
- Knowledge and data mapping rules are onerous to bootstrap but lightweight to maintain.

  => Lots of initial testing datasets help us foresee data heterogeneity.