AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

March 15, 2018

Nicolas Paris, MSc.[1,2,3], Michael Mendis[4], Christel Daniel, MD, Ph.D[1,3], Shawn Murphy, MD, Ph.D[4], Xavier Tannier, Ph.D[3,5] and Pierre Zweigenbaum, Ph.D[2]

S43: Research Informatics Infrastructure

i2b2 implemented over SMART-on-FHIR

[1]WIND-DSI, AP-HP, Paris, France;
[2]LIMSI, CNRS, Université Paris-Saclay, Orsay, France;
[3]INSERM, UMR_S 1142, LIMICS, Paris, France;
[4]Partners Healthcare In., Boston, MA, USA;
[5]Sorbonne Universités, UPMC Univ Paris 06, France

Cohort Discovery

- Subsetting patients based on eligibility criteria
- Key step in medical research
- Observational studies & Clinical trials
- At national or international scale

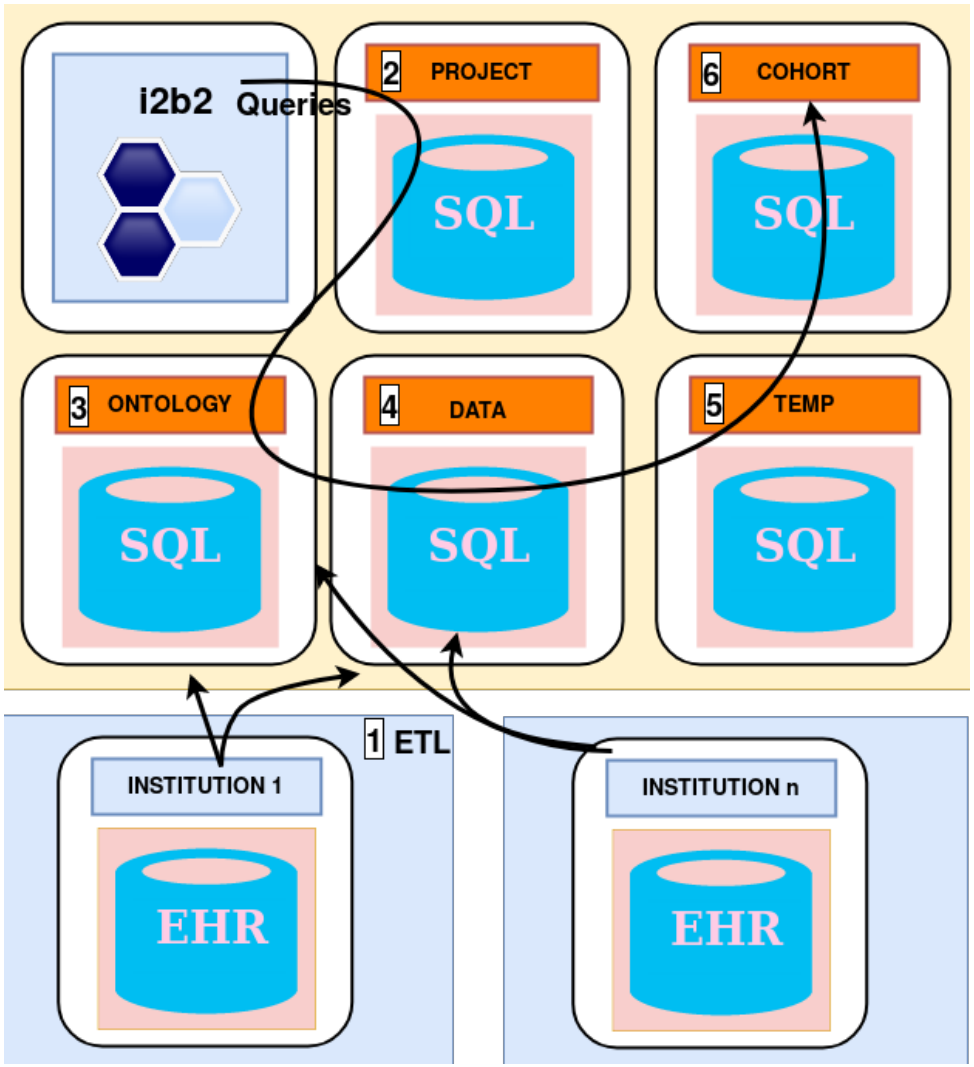
$\rightarrow$ Fast & Precise cohort discovery is crucial

Cohort Discovery Tools

- i2b2 is leader in cohort discovery
- Federating i2b2 instances (Shrine, Insite, TrinetX. . .)
- Analytic tools extend i2b2 (transmart) to handle genomic
- Altogether provides a 360°view of patients and hospitals
- About 200 implementations worldwide

$\rightarrow$ Large scale cohort discovery needs i2b2 democratisation

i2b2 Architecture



The star schema:

- Cohort selection is fast
- Adapts to local needs
- Simple to load

The ontology design:

- Stores any terminology
- Represents information easily
- Customizable, extensible concepts

$\rightarrow$ Flexible, Efficient, Simple.

Improving i2b2

Data Veracity Challenge

- ETL complexity may affect data consistency
- ETL latency compromizes data currency

Data Variety Challenge

- Data variety needs technology variety (text, time series, imaging. . .)
- Ontology variety needs terminology mapping

Data Volume Challenge

- Duplicating big-data = big-data x 2
- Analyzing big-data needs specialized technologies

Fast Healthcare Interoperability Resources

HL7 FHIR

- FHIR is an Agile and RESTful Approach to Healthcare Information Exchange
- FHIR covers most patient & healthcare-related concepts
- FHIR specifies both a *Search* & a *Terminology Mapping* API
- Open-Source community massively adopts FHIR
- $\approx$ 50 SMART-on-FHIR apps !

Bridging i2b2 & FHIR

Related work:

- Mobile phone pushing data into i2b2 [1]
- i2b2 as a FHIR endpoint [2]
- SMART-on-FHIR applications running on top of i2b2 [3]

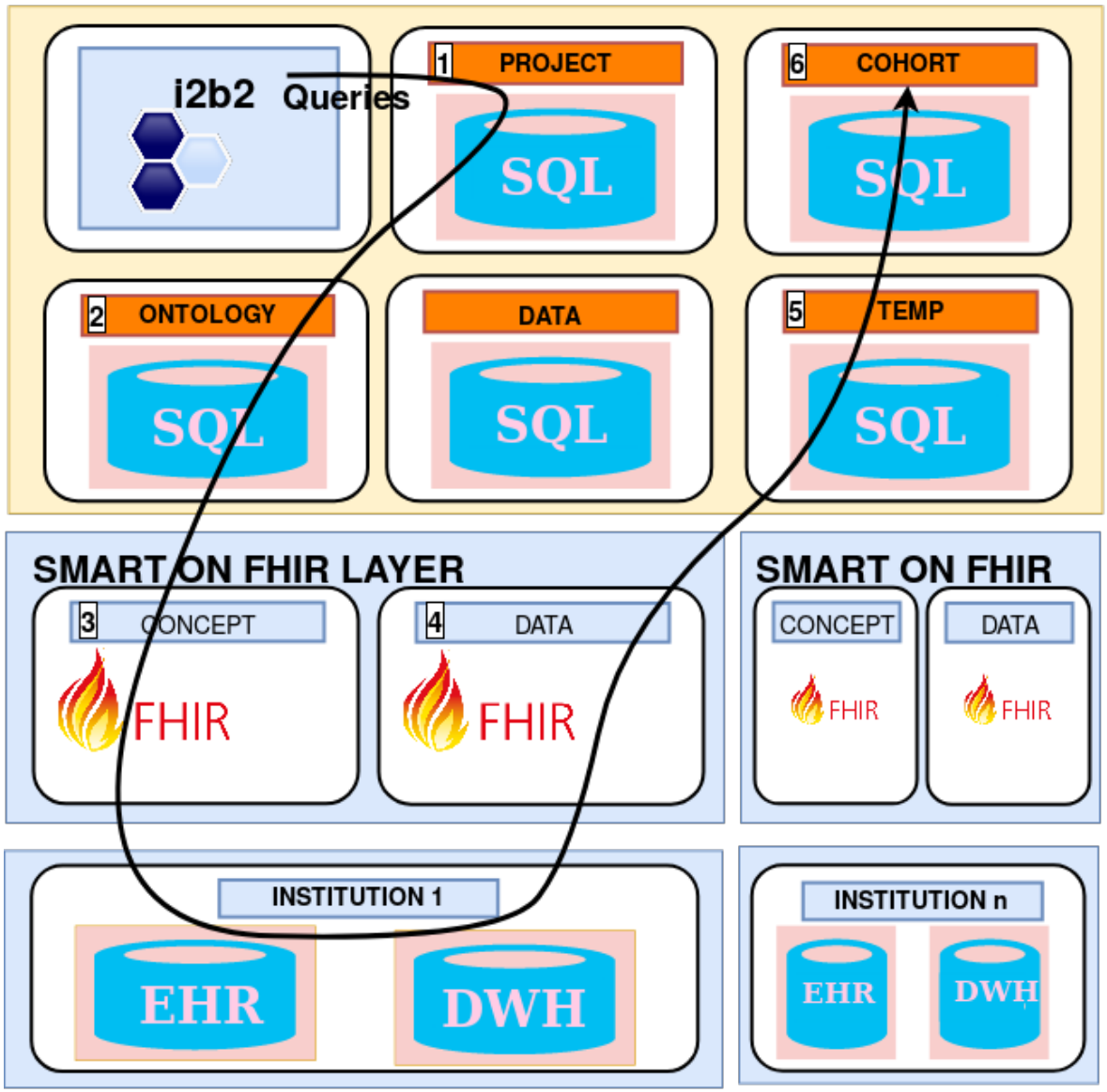
$\rightarrow$ Present work = an i2b2 connector to FHIR endpoints

[1] Pfiffner et al. "C3-PRO: Connecting researchkit to the health system using i2b2 and FHIR", Mar 2016.

[2] Boussadi et al. "FHIR layer implemented over i2b2", Dec 2017

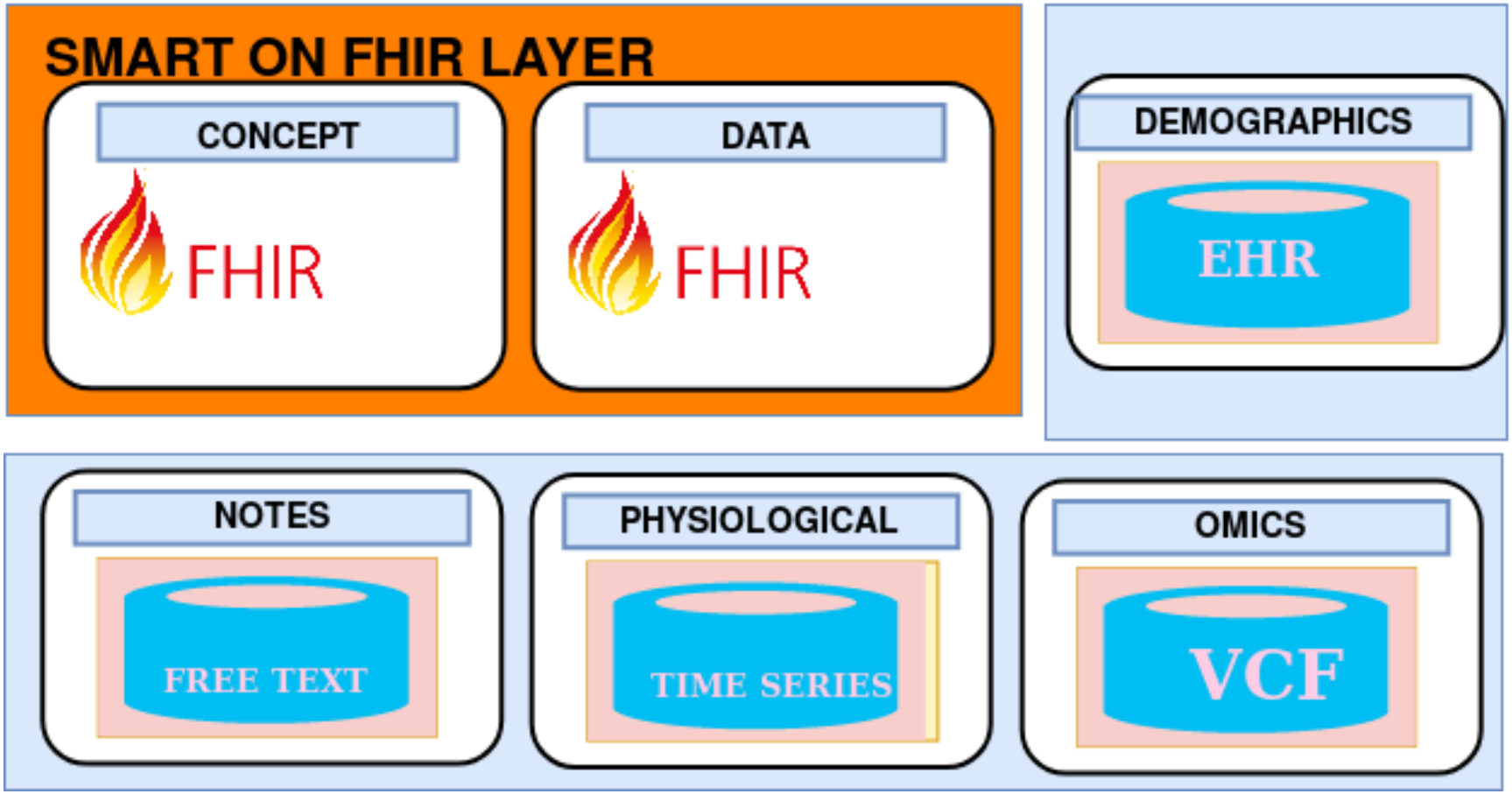
[3] Wagholikar et al. "SMART-on-FHIR implemented over i2b2", Jun 2016

Overall i2b2-FHIR Architecture



FHIR Endpoint Implementation

- data variety ↔ database variety



Data Search

- Build on the FHIR Search Specifications
- GET

  `<FHIR-API>/<Resource>?elements=<elements>&code=<codes>`
  `&date=gt<date inf>&date=lt<date sup>&<custom filter>`

- RETURNS

  Patient Set:

```
{ "resourceType": "Patient", "id": "8989" }
```

  Encounter Set:

```
{ "resourceType": "Encounter", "id": "100100", "patient": { "
    reference": "Patient/001"} }
```

  Temporal Set:

```
{ "resourceType": "Observation", "subject":{"reference":"Patient/0
    01"}", effectiveDateTime": "2018-03-15T10:30:00" }
```

Connector Configuration

- One YAML configuration file per endpoint

```
version: dstu3
Patient:
    patientUriPath: $.resource.id
    patientUriField: id
Observation:
  - patientUriPath: $.resource.subject.reference
  - encounterUriPath: $.resource.context.reference
  - instanceUriPath: $.resource.id
  - datePath: $.resource.effectiveDateTime
  - patientUriField: subject
  - encounterUriField: context
  - instanceUriField: id
  - dateField: effective
```

- Institutions can use FHIR Extensions

Concept Mapping

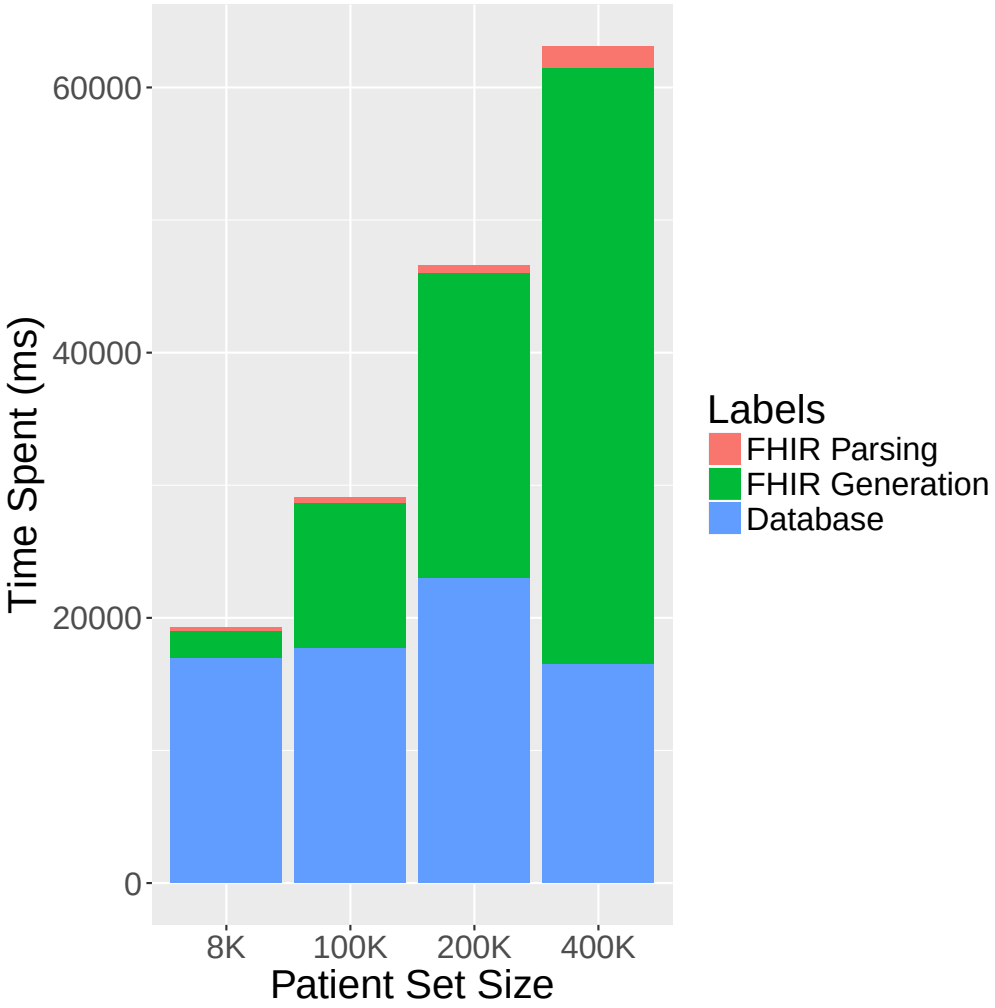
- Takes advantage of FHIR standard Concept Mapping
- GET

  `<FHIR-API>/ConceptMap`
  `?source-code=<codes>&source-system:in=<code-system>`
- RETURNS

```
{ "resourceType": "ConceptMap",
  "group": [{
    "source": "http://hl7.org/fhir/address-use",
    "target": "http://hl7.org/fhir/v3/AddressUse",
       "element": [{
          "code": "home",
          "target": [
             { "code": "H", "equivalence": "equal"},
             { "code": "house", "equivalence": "equal"}
             ]}],
        }]
}
```

- Filters Equivalent codes

Benchmarks



5B physiological table

- Bottleneck = json generation
- Linear performances
- Under the minute

Results

Data Veracity Challenge

- ETL are not mandatory anymore
- Some queries can be real-time

Data Variety Challenge

- Concept Mapping $\leftrightarrow$ standard terminologies
- Dedicated DBs can exploit heterogeneous data types
- Multiple center can be queryied together

Data Volume Challenge

- Query data in place without duplication
- Allows big-data handling

Discussion

Limitations:

- EHR vendors not all FHIR compliants
- Concept mapping remains to be done
- FHIR has no asynchronous mode

Perspectives:

- FHIR GraphQL & Bulk Export would improve
- Multiple FHIR endpoint implementation
- Oauth 2.0 securisation implementation
- Release as a SMART-on-FHIR app

Conclusion

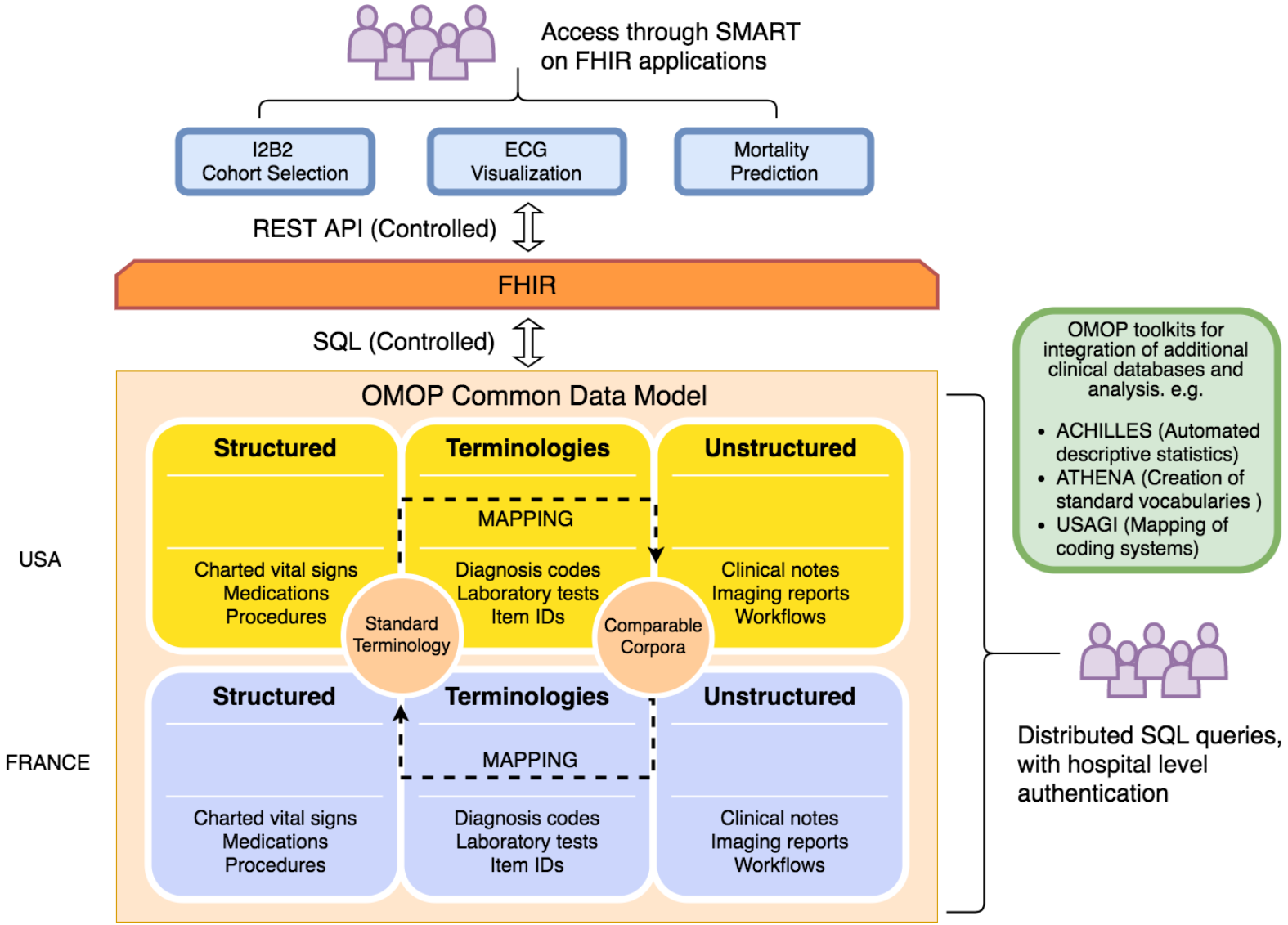
This work opens new areas:

- Discover cohort over countries
- Explore new data types

Future directions:

- Semi-Automatic Concept Mapping methodology
- Create FHIR-OMOP connectors

Overall Project



Questions ?

nicolas.paris@aphp.fr