



# InteropEHRate

**D5.18**

## **Library and tool for data mapping and conversion - v1**

**ABSTRACT**

This document accompanies the release of the software components that are responsible for the conversion of Electronic Health Records across local and national healthcare standards. It covers the progress made on the definition and implementation of health knowledge for knowledge-based data integration, as well as the supporting libraries and tools.

<b>Delivery Date</b>	January 22 <sup>nd</sup> , 2020
<b>Work Package</b>	WP5
<b>Task</b>	T5.5
<b>Dissemination Level</b>	Public
<b>Type of Deliverable</b>	Demonstrator
<b>Lead partner</b>	UniTN



This document has been produced in the context of the InteropEHRate Project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826106. All information provided in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose.



This work by Parties of the InteropEHRate Consortium is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

DRAFT

## CONTRIBUTORS

	Name	Partner
Contributors	Gábor Bella	UNITN
Contributors	Simone Bocca	UNITN
Contributors	Alessio Graziani	ENG
Contributors	Stefano Dalmiani	FTGM

## LOGTABLE

Version	Date	Change	Author	Partner
0.95	09-01-19	Document created	Gábor BELLA, Simone BOCCA	UNITN
1.0	14-01-20	Quality Review	Argyro Mavrogiorgou	UPRC
1.1	21-01-20	Fixed document structure to be more conform to template	Gábor BELLA, Simone BOCCA	UNITN
Vfinal	22-01-20	Final check and version for submission	Laura Pucci	ENG

## ACRONYMS

Acronym	Term and definition
EHR	Electronic Health Record
FHIR	Fast Healthcare Interoperability Resources
ICD	International Classification of Diseases
LOINC	Logical Observation Identifiers Names and Codes
SNOMED CT	Systematized Nomenclature of Medicine -- Clinical Terms
WHO ATC	World Health Organization's for active substances of pharmaceutical products
CDA	Clinical Document Architecture
SEHR	Smart Electronic Health Record

DRAFT

TABLE OF CONTENTS

1. INTRODUCTION .....1

1.1 Scope of the document .....1

1.2 Intended audience.....1

1.3 Structure of the document .....1

1.4 Updates with respect to previous version (if any) .....1

2. SW DESCRIPTION .....2

2.1 Pivot Knowledge .....2

2.2 Mappings to Pivot Knowledge.....3

3. OVERVIEW .....4

3.1 Pivot Knowledge .....5

3.1.1. Building guide .....5

3.1.2. Installation guide .....6

3.1.3. User guide.....6

3.2 Mappings to Pivot Knowledge.....6

3.2.1 Building guide .....7

3.2.2 Installation guide .....7

3.2.3 User’s guide.....7

LIST OF TABLES

- Table 1 - SNOMED Standard
- Table 2 - ICD-10 Standard
- Table 3 - FHIR Standard
- Table 4 - LOINC Standard
- Table 5 - FHIR Ontology Model
- Table 6 - ICD-10 Standard
- Table 7 - CDA to FHIR Data Mapping Model
- Table 8 - Implementation status of conversion services and tools
- Table 9 - Pivot knowledge implementation status
- Table 10 - Mapping implementation status

# 1. INTRODUCTION

## 1.1 Scope of the document

This document provides details about the current (at the time of writing the deliverable, end of 2019) demonstrator release of data, tools, and libraries for health data mapping and conversion. The details refer to:

- an overview of the conversion tasks addressed so far;
- the data files released or in progress;
- how to build, install, use the data files using Platform tools;
- release modalities.

## 1.2 Intended audience

The document is intended to all people interested to have an overview of the progress made on cross-border conversion of EHR data.

## 1.3 Structure of the document

Knowledge components related to conversion are grouped into *pivot knowledge* and *mapping knowledge*. Accordingly, section 2 contains two sub-sections describing the knowledge packages corresponding to each group. Likewise, section 3 provides a longer textual overview of these components, as well as building, installation, and succinct user guides.

## 1.4 Updates with respect to previous version (if any)

Not applicable, this is the first version of the deliverable.

## 2. SW DESCRIPTION

### 2.1 Pivot Knowledge

STANDARD NAME	SNOMED CT International
IMPORT FILE FORMAT	Excel (xls)
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Import Files/SNOMED)

*Table 1 - SNOMED Standard*

STANDARD NAME	ICD-10
IMPORT FILE FORMAT	CSV
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Import Files/ICD)

*Table 2 - ICD-10 Standard*

STANDARD NAME	FHIR 4.0 IPS
IMPORT FILE FORMAT	Excel (xls)
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Import Files/FHIR)

*Table 3 - FHIR Standard*

STANDARD NAME	LOINC
IMPORT FILE FORMAT	Excel (xls) (zip)
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Import Files/LOINC)

*Table 4 - LOINC Standard*

MAPPING NAME	FHIR Formal Ontology Model
IMPORT FILE FORMAT	OWL
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Types/iehr-owl-09012020.owl)

*Table 5 - FHIR Ontology Model*

## 2.2 Mappings to Pivot Knowledge

STANDARD NAME	ICD-9-to-10 mappings
IMPORT FILE FORMAT	CSV
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/knowledge-mgmt-tool.git</a> (Import Files/ICD)

*Table 6 - ICD-10 Standard*

MODEL NAME	CDA to FHIR Data Mapping Model
MODEL FILE FORMAT	TTL
REPOSITORY	<a href="http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/data-mapping-tool.git">http://iehrgitlab.ds.unipi.gr/interopehrate/health-tools/data-mapping-tool.git</a> (Models/cda_result.xml-model_final.ttl)

*Table 7 - CDA to FHIR Data Mapping Model*

DRAFT



### 3. OVERVIEW

The InteropEHRate Platform includes tools that adopts a knowledge-based approach to automating data integration and interoperability. In practice, this means that the majority of the implementation efforts go into specifying the formal knowledge that enables conversions, instead of procedural solutions (libraries or tools).

As detailed in D5.9 [\[4\]](#), mapping and conversion apply to the following kinds of health information:

- EHR data schemas;
- EHR data values, including:
  - coded values referring to healthcare concepts,
  - natural language text,
  - other kinds of data values that need to be transformed.

We call the conversion of natural language text across languages “translation”, which is addressed in a different deliverable and is thus outside the scope of this one.

The knowledge-based automation of mapping and conversion methods adopted by the InteropEHRate project require that health knowledge allowing such conversions should be formalised and integrated into the InteropEHRate Health Services. This includes:

- pivot representations as defined by the Interoperability Profile [\[2\]](#):
  - FHIR data schemas,
  - ICD-10 codes for diseases,
  - LOINC codes,
  - SNOMED CT codes,
  - WHO ATC codes for pharmaceutical products;
- mappings between local representations and pivot representations for all of the above, to be implemented for the four participating hospitals from four countries (Italy, Belgium, Romania, Greece).

Beyond the formalisation and integration of knowledge and mappings, the project also provides to hospitals high-level data conversion services, as well as interactive tools that let data scientists (employed by the hospitals) specify the mapping procedures and the corresponding knowledge.

The tools are released as part of the *InteropEHRate Health Tools* macro-component, itself described within the Platform deliverable [\[3\]](#). The table below provides the list of services and tools, specified to be implemented.

Requirement	Software Component	Status
Interactive definition of pivot knowledge and terminology mappings	Knowledge Modeller Tool	V1 done, deployed as part of D5.16 (Platform)
Programmatic (batch) integration of massive amounts of pivot knowledge and terminology mappings	Knowledge Integration component	V1 done, deployed as part of D5.16 (Platform)
Interactive and programmatic (batch) conversion of local EHRs to the FHIR-based SEHR format	Data Mapping Tool	V1 done, deployed as part of D5.16 (Platform)

Service to initiate the automated conversion of an EHR	InteropEHRate Conversion Services	by end 2020
Service to retrieve an already converted EHR in serialised FHIR format	InteropEHRate Conversion Services	by end 2020
Low-level service to convert individual terms across standards	InteropEHRate Conversion Services	already exists on the Platform level, full implementation by end 2020

*Table 8 - Implementation status of conversion services and tools*

### 3.1 Pivot Knowledge

Pivot knowledge consists of international healthcare standards that are used by the Interoperability Profile [2]. By “implementation” we mean:

- conversion of these standards from the formats in which they are distributed to the formats expected by the InteropEHRate Data Integration Platform;
- archival of such converted resources for future reuse, including documentation;
- ingestion of the standards into working instances of the InteropEHRate Health Services.

Furthermore, whenever available, pivot knowledge needs to be implemented in five languages, corresponding to the four hospitals participating in the InteropEHRate project (Italian, French, Romanian, Greek) plus English as a lingua franca.

The table below shows the progress made so far on implementing pivot knowledge until the end of 2019.

Standard	Use	Size	Done	Status	Language support
SNOMED CT International	reference terminology	1.5 million labels	YES	full 2017 version imported	English
ICD-10	disease codes	69,000 labels	YES	WHO version imported	English, Italian
LOINC	lab observations	91,388 labels	YES	fully imported	English
WHO ATC	drugs	>13476 labels	NO	to be done	
FHIR 4.0 IPS	EHR schemas for International Patient Summary	458 attributes	YES	90% coverage	English

*Table 9 - Pivot knowledge implementation status*

#### 3.1.1. Building guide

The original knowledge files provided by third-party distributors need to be:

1. converted to the knowledge representation format accepted by the Platform;
2. uploaded into the platform.

Conversion is executed using the appropriate Python script released as part of the Platform deliverable [\[D5.16\]](#), in the following way:

```
python script-name.py
```

where *script-name* refers to the script corresponding to the type of knowledge being converted (e.g. LOINC or SNOMED). The output of the conversion is one or more Excel files containing knowledge in its final uploadable form.

### 3.1.2. Installation guide

Once the knowledge files are correctly converted, a command line tool is in charge of uploading them to the knowledge base. The following command starts the knowledge import procedure:

```
updateuk -i import-file.xls
```

Where import-file.xls is an Excel file containing the knowledge which is to be imported.

### 3.1.3. User guide

The knowledge integrated into the Platform can be used:

- through the Knowledge Viewer tool for browsing and understanding;
- automatically by the InteropEHRate Conversion Services to convert EHRs to the pivot SEHR representation;
- through the Data Mapper Tool where the user (the Data Scientist) checks that the data structure and values are correctly mapped to the pivot representations.

The Knowledge Viewer and Data Mapper tools will be provided separate user’s guides as part of the deliverable of their final versions.

## 3.2 Mappings to Pivot Knowledge

Local EHRs and the terminology used within have so far been analysed on an example EHR IPS file provided by FTGM, Italy. This local EHR is structured along the XML-based CDA standard. It uses codes from ICD-9, WHO ATC, and LOINC. Given that ATC and LOINC are also part of the Interoperability Profile, terminological mappings were only necessary between ICD-9 and ICD-10. In addition, the CDA format needed to be mapped to FHIR.

The EHR formats and related knowledge of the remaining three hospitals are yet to be implemented.

Local Standard	Local standard user	Pivot Standard	Implemented	Status	Language support
ICD-9	FTGM (Italy)	ICD-10	YES	only 1-to-1 equivalence mappings that are covered by ICD-10	From Italian towards English
Patient Summary XML CDA	FTGM (Italy)	FHIR 4.0 IPS	YES	mapping only covers the contents of the example file	From Italian towards English

				provided by FTGM	
SumEHR Patient Summary	CHU (Belgium)	FHIR 4.0 IPS	NO	to be done	

*Table 10 - Mapping implementation status*

### 3.2.1 Building guide

Building (converting) mappings is performed exactly the same way as for pivot knowledge, see section 2.2.1.

### 3.2.2 Installation guide

Installation (importing) of mappings is performed exactly the same way as for pivot knowledge, see section 2.2.2.

### 3.2.3 User's guide

The browsing and use of mappings are performed exactly the same way as for pivot knowledge, see section 2.2.3.

DRAFT

## REFERENCES

- [1] **[D2.1]** InteropEHRate Consortium, User Requirements for cross-border HR integration - V1, 2019.  
<https://www.interopehrate.eu/resources/>
- [2] **[D2.7]** InteropEHRate Consortium, FHIR Profile for EHR interoperability - V1, 2019.  
<https://www.interopehrate.eu/resources/>
- [3] **[D5.7]** InteropEHRate Consortium, Design of the Health Data Integration Platform - V1, 2019.
- [4] **[D5.9]**: InteropEHRate Consortium, Design of the Data Mapper and Converter to FHIR - V1, 2019.
- [5] **[D5.16]**: InteropEHRate Consortium, Data integration platform for healthcare professionals – v1, 2019.

DRAFT