The Eleventh International Conference on Global Health Challenges
GLOBAL HEALTH 2022

Challenges and Advanced Technology Solutions in Applied Health Informatics
(Case studies from the Bulgarian eHealth National Scientific Program)

Evgeniy Krastev[1*], Dimitar Tcharaktchiev[2*]
Petko Kovachev[1], Simeon Abanos[1]

[1]Faculty of Mathematics and Informatics, Sofia University St. Kliment Ohridski, **Bulgaria**
[2]Medical University Sofia, **Bulgaria**
[*] **Presenter**

Professor Evgeniy Krastev

Evgeniy Krastev Is a Professor the department of Mechatronics, Robotics and Mechanics at Sofia University where in 1986 he received PhD. degree in Mathematics and Computer science. He is Head of the M.Sc. Degree program Bio-Medical Informatics at the Faculty of mathematics and informatics. He is team leader of the National Scientific Program eHealth in Bulgaria He has delivered lectures at SUNY-State University of New York (Empire State College), (USA), Frederick Institute of Technology(Cyprus), at American University of Technology and at American University of Science and Technology (Lebanon). The professional interests of Professor Krastev include Object oriented software development, Health informatics, Kinematic control and software development for robot arms, Modeling, analysis and simulation of Business processes and Business rules, Relational database management and programming, Modern operating systems.

Professor Dimitar Tcharaktchiev

Dimitar Tcharaktchiev is Professor at Medical University– Sofia and Head of the Department of Medical Informatics of the University Specialized Hospital of Active Treatment of Endocrinology. He holds Ph.D. in Medical Informatics. He is member of the Bulgarian Institute for Standardization–BDS, Bulgarian representative (1998 – 2002) in the European Federation of Medical Informatics (EFMI) and in the European Organization of Standardization–TC(Technical Committee) for Medical Informatics – CEN TC251 (from 2015). Invited scientist (2001-2004) to the University of Paris VI, the European University Hospital "Georges Pompidou", member of the team which won the eHealth Prize in 2003 of the European Commission, winner of the "Rolf Hansen" memorial award (2011) of the European Federation of Medical Informatics. Prof. Tcharaktchiev is chairman of Association PROREC Bulgaria, Member of EUROREC Institute and Leader of the National Scientific Program eHealth in Bulgaria.

Case studies

1. International exchange of ePrescriptions
   - 1.1 Methods
   - 1.2 Results
   - 1.3 Summary
2. International Patient Summary (IPS)
   - 2.1 Methods
   - 2.2 Results
   - 2.3 Summary
3. OMOP Common Data Model Discussion
   - 3.1 Methods
   - 3.2 Results
   - 3.3 Summary
4. Discussion

1. International exchange of ePrescriptions

ePrescription service

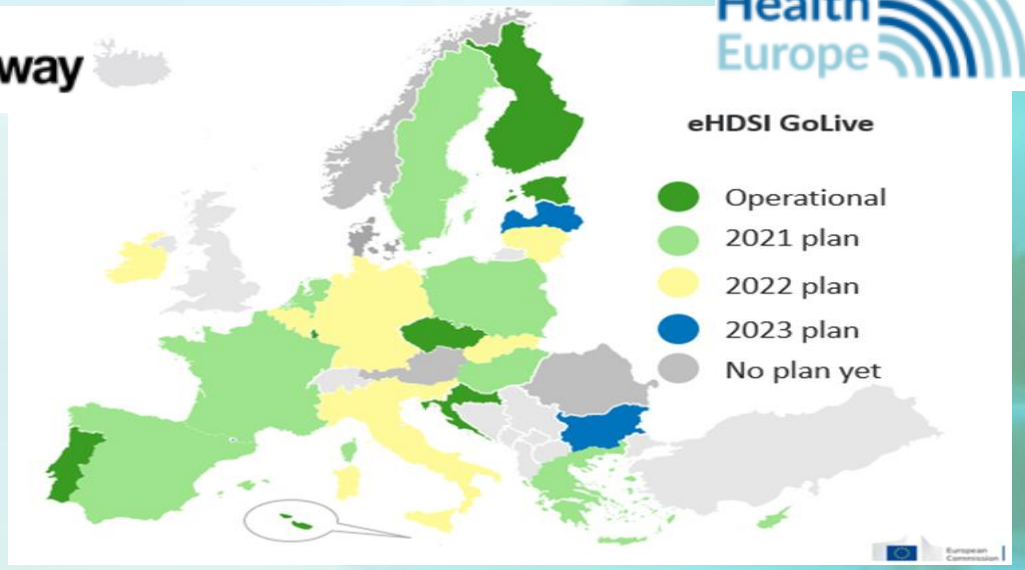
Canada Health Infoway

PrescribeIT
CANADA'S ELECTRONIC PRESCRIPTION SERVICE

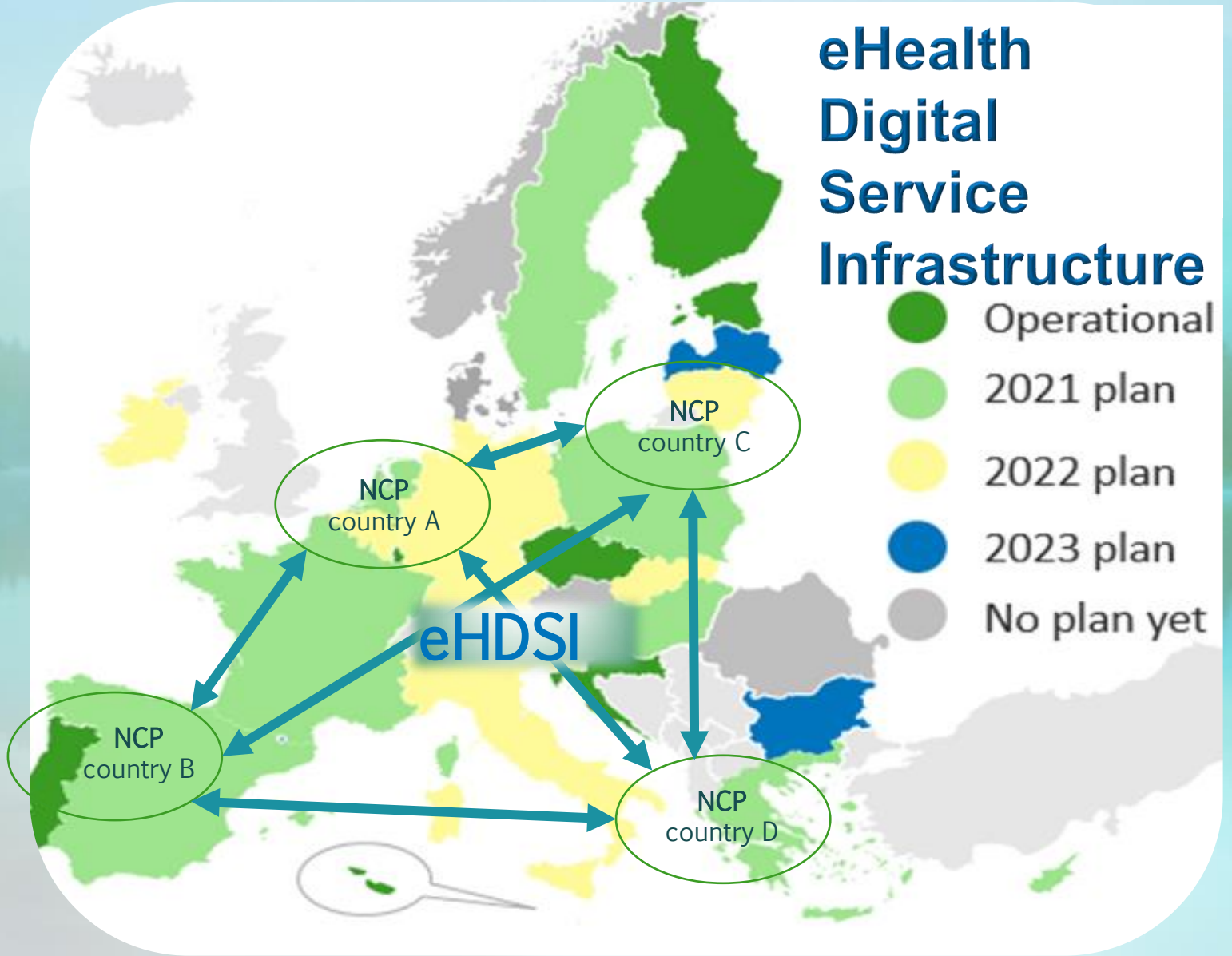
surescripts

eRx
script exchange

Digital Health Europe



1. International exchange of ePrescriptions



1. International exchange of ePrescriptions

eHealth Digital Service Infrastructure (**eHDSI**):

- ✓ Provides a **common infrastructure** for cross-border. exchange of ePrescriptions (**eP**) at EU level.
- ✓ Establishes a **common ePrescription dataset.**
- ✓ Determines **use cases and functional requirements.**
- ✓ Establishes National Contact Points (**NCP**).

Challenges:

- ❑ eP deployment in a single country.
- ❑ eP cross- border exchange between multiple EU counties (**27**) with different Clinical Information Models(**CIM**) and legislation.

Research Objective:

- ✓ Propose an approach for cross-border exchange of eP content
- ✓ Provide uniform representation of the eHDSI basic dataset of eP by means of QR code

1.1 International exchange of ePrescriptions- Methods

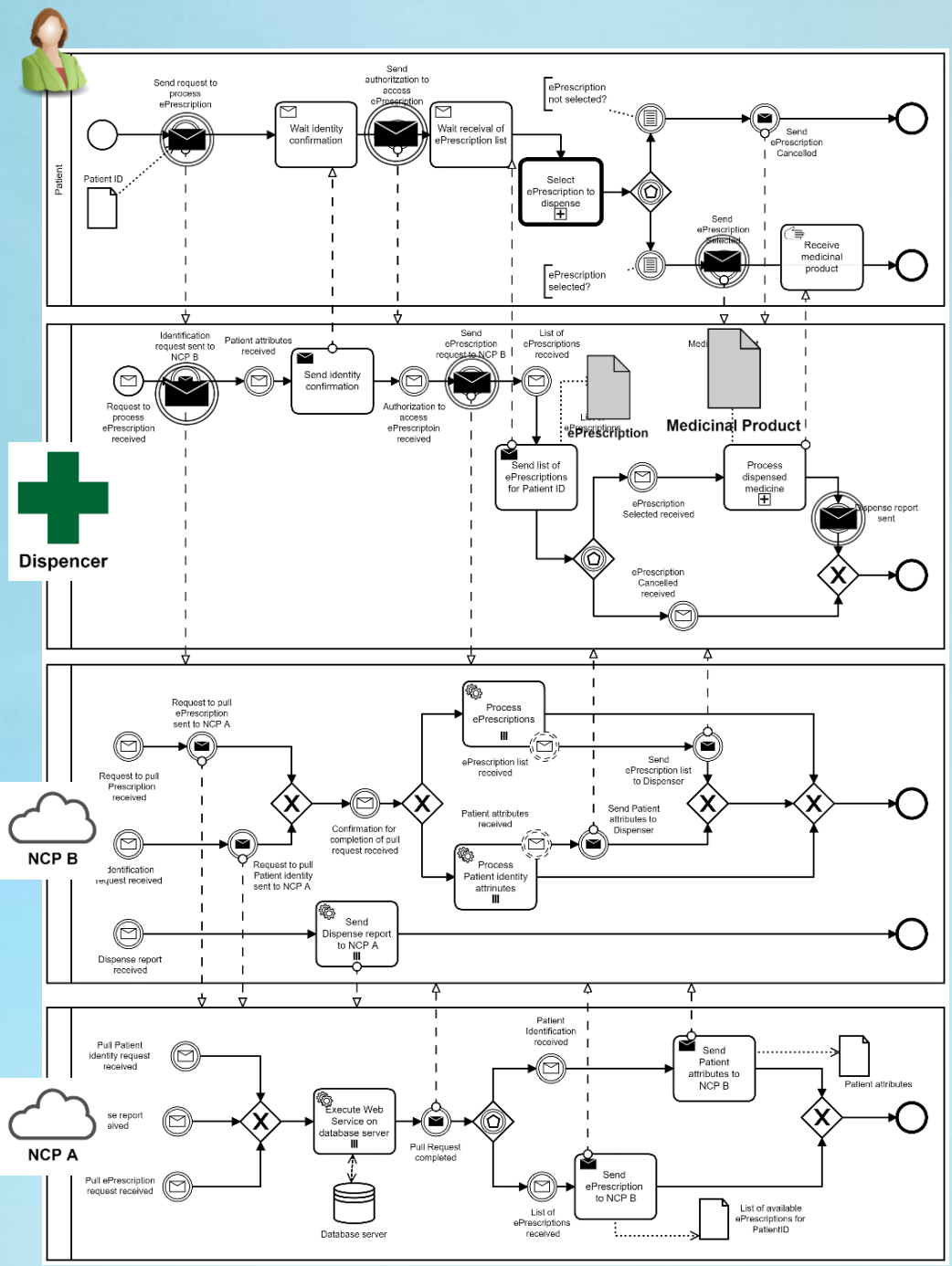
Primary Use case for eP cross-border exchange:

Actors:

- **Patient**
- **Dispenser**(Dispenser Provider)
- National Contact Points (**NCP**) in country A (patient's country)
- National Contact Points (**NCP**) in country B

Message exchange

- Assumes **permanent network availability** between pharmacies and NCPs
- Involves **large number of bidirectional transformations** of eP data between different CIM
- **Centralized**, NCP- based flow of events
- Allows **merely technical interoperability**



Centralized, NCP-based flow of events

1. Patient identification
2. Patient authorizes eP extract
3. Dispencer extracts Patient's ePs
4. Patient selects Product

5a. Patient gets Medicinal Product

5b. NCP A saves Dispense report

Business Process Model & Notation (BPMN) diagram (OMG standard) of the primary Use case for cross-border exchange of EPrescriptions

1.1 International exchange of ePrescriptions- Methods

Inversion of Control:

- **The *Patient- based centralized control* over the ePrescription processing**
- The eHDSI basic data set of eP is represented in a **CIM- independent XML model**
- Represent the eP dataset in **QR code technology**

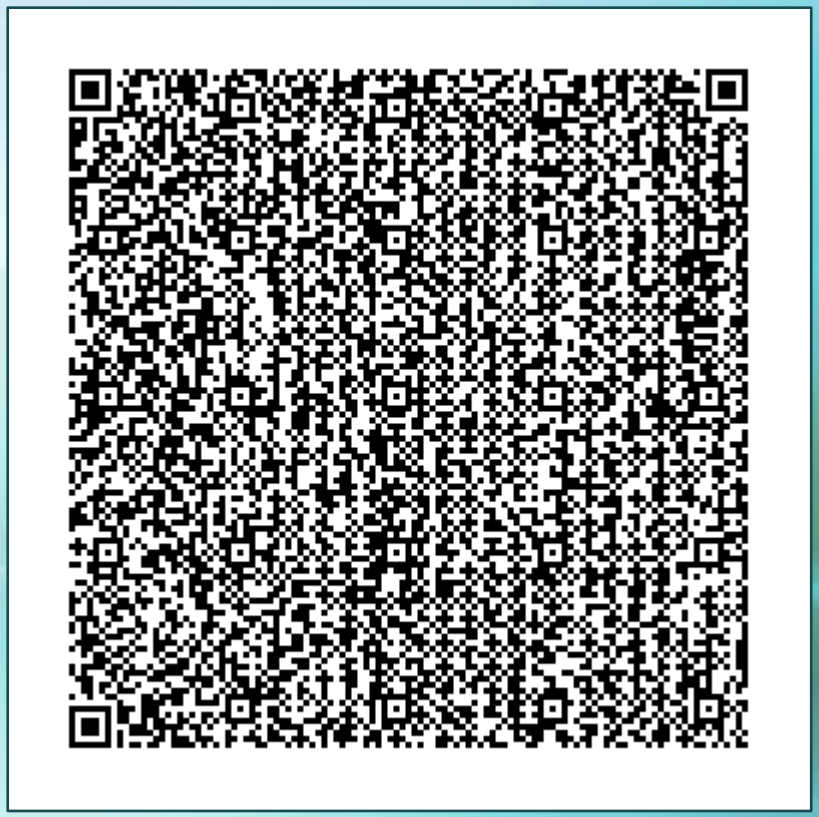
Expected benefits:

- **QR code storage** of eP details is **aligned** with recently introduced **regulations for identification of medicinal products**
- **Seamless integration** of the eP service in pharmacy software

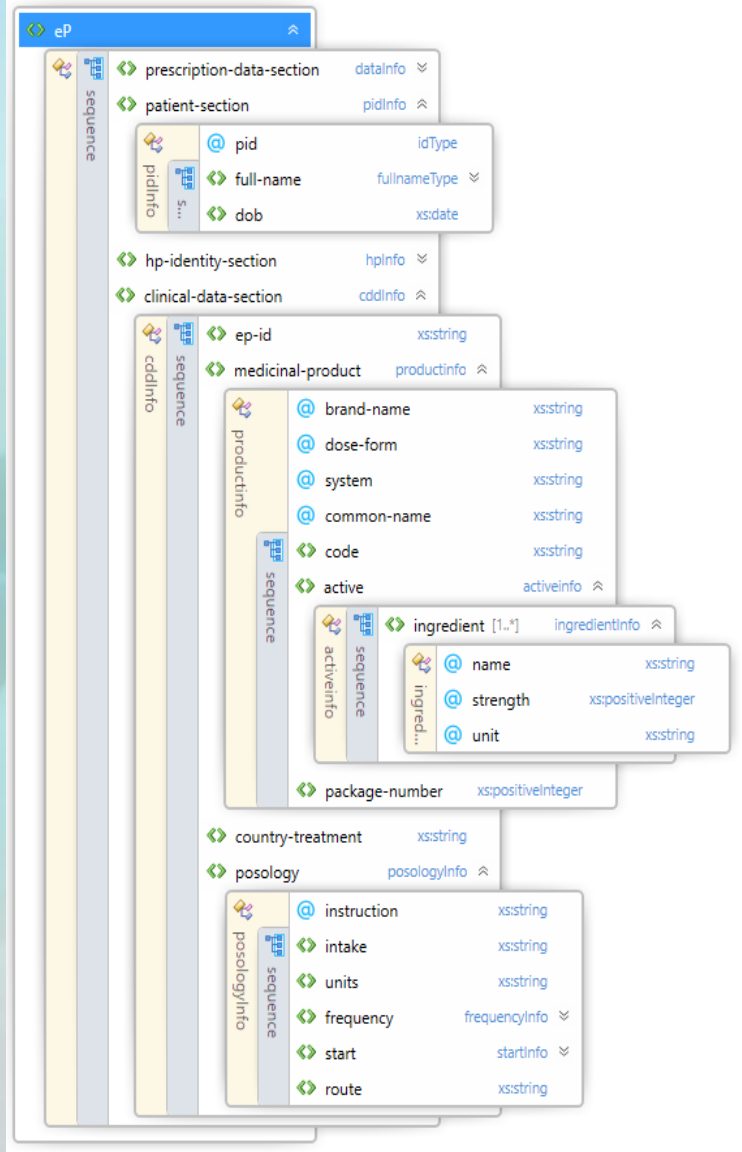
1.2 International exchange of ePrescriptions- Results

Computer experiments for generating ePrescription content:

- A CIM-independent XML model is developed
- Web application to generate eHDSI “friendly” real-life data for ePrescription
- The ePrescription content is stored in instances of that model
- The XML is compacted in JSON and stored in QR code



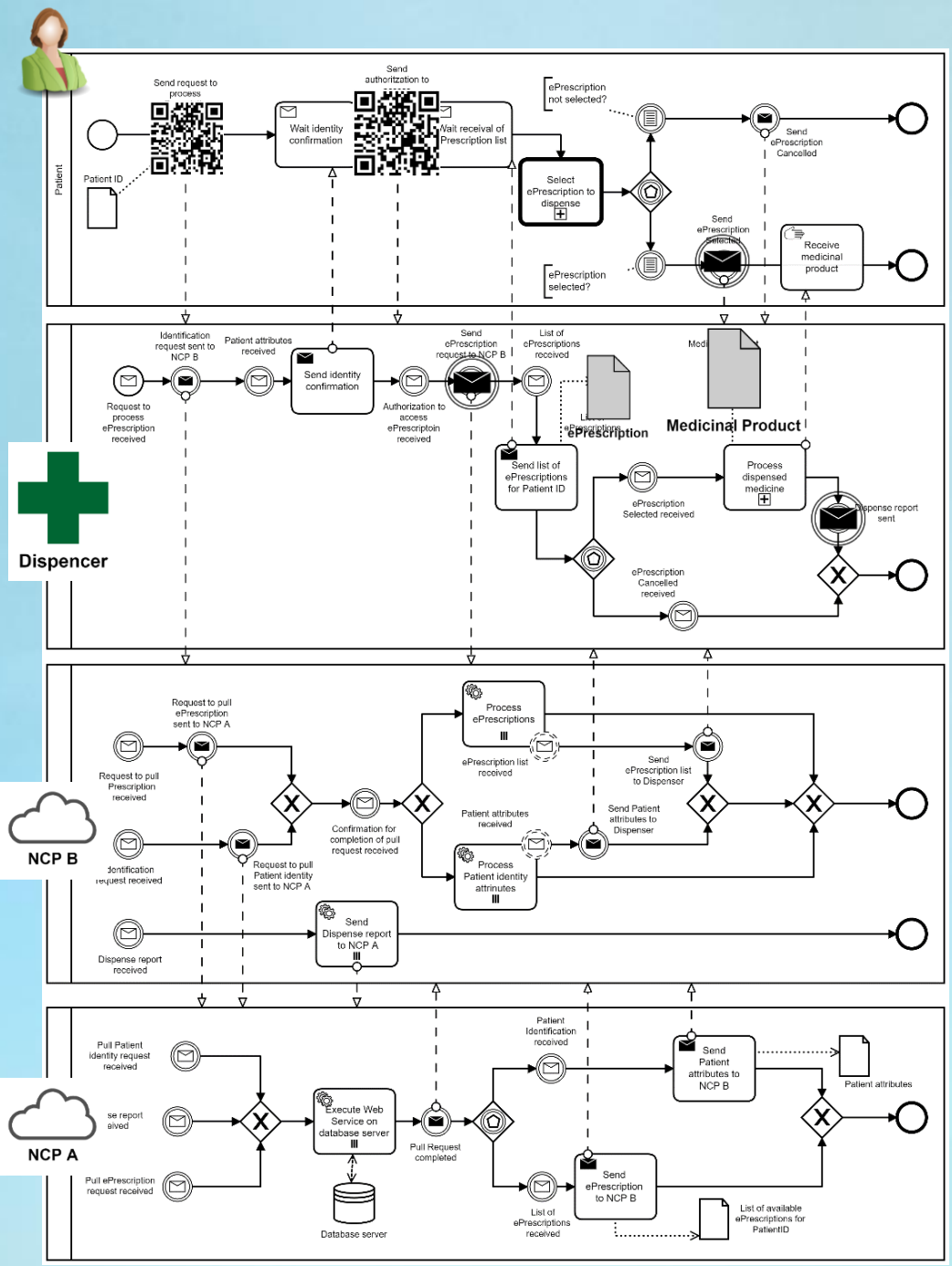
1.2 International exchange of ePrescriptions- Results



XML scheme of the eHSDI basic dataset of an ePrescription.

Main sections:

1. **Patient identification**
2. **Health Professional identification**
3. **Prescription clinical data**
4. **Prescription data**



Patient-based flow of events

1a. Patient identification

1b. Patient authorizes eP extract

2. Dispenser serves the product

3a. Patient receives the Product

3b. NCP A saves Dispense report

Note:

1.The ePrescription stores in QR code data for a single Medicinal product.

2. QR code technology provides means for Patient identification.

1.2 International exchange of ePrescriptions- Results

Inversion of Control reproduces the natural way to handle a prescription, where the **Patient** has full control over management of prescription data.

Computer experiments:

- ✓ **Demonstrate practicability** of the proposed approach
- ✓ Allow significant reduction of message exchanges
- ✓ Enable off-line **feasibility of eHDSI functional requirements**
- ✓ **Based on QR code representation** of ePrescription content
- ✓ Allow **seamless integration** in existing pharmacy software
- ✓ **Complies with regulations** for unique identification of medicinal products

1.3 International exchange of ePrescriptions- Summary

Advantages:

- ❑ **Satisfies** the **eHDSI functional requirements**
- ❑ Significantly **reduces** message exchanges
- ❑ Enables **off-line operation** of the ePrescription service
- ❑ Employs **XML scheme** that is **independent** of any CIM
- ❑ ePrescription cross-border **exchange** between **different CIM**

Disadvantage

- ✓ Needs support by EU governments and service providers

Future research work

- ✓ Develop algorithms for secure processing of ePrescription content in QR code
- ✓ Improve the process of compiling ePrescription content

Video presentation in YouTube

P. Kovachev, "Uniform representation of ePrescription content in QR code," YouTube, 2021. [Online].

1.3 International exchange of ePrescriptions- Summary

Publications

Evgeniy Krastev, Petko Kovachevachev, Dimitar Tcharaktchiev, Simeon Abanos, **Using QR Code for Uniform Representation of Content in Cross-border Exchange of ePrescriptions in the EU**, 31th Medical Informatics Europe conference (MIE 2021), May 29 – 31 (2021), Athens, Greece, pp. 684 – 68. DOI:10.3233/SHTI210259,

P. Kovachev and E. Krastev, "**Modelling and Management of ePrescriptions on openEHR Platform in Bulgarian eHealth**", in The Ninth International Conference on Global Health Challenges (25- 29 October), Nice, France, 2020. Available:. [retrieved: October, 2020]. ISSN: 2308-4553 (ThinkMind Digital Library open access)

2. International Patient Summary (IPS) Case study introduction

The International Patient Summary (IPS) has a long story:

- **epSOS**, meaning "Smart Open Services for European Patients"- cross- border pilot projects for Patient Summary and ePrescriptic exchange
- Joint Action to support the eHealth Network project (**JAseHN**)
- eHealth Digital Service Infrastructure (**eHDSI**)

eHealth DSI Deploying Countries	PS	eP
Austria	♥	♥
Croatia	♥	♥
Cyprus	♥	♥
Czech Republic	♥	
Estonia	♥	♥
Finland		♥
France	♥	
Germany	♥	
Greece	♥	♥
Hungary	♥	♥
Ireland	♥	♥
Italy	♥	♥
Luxembourg	♥	
Malta	♥	
Portugal	♥	♥
Sweden		♥
Switzerland	♥	♥

epSOS

JASEHN
Joint Action to Support the eHealth Network

International Patient Summary (IPS) for Unscheduled, Cross-border Care

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

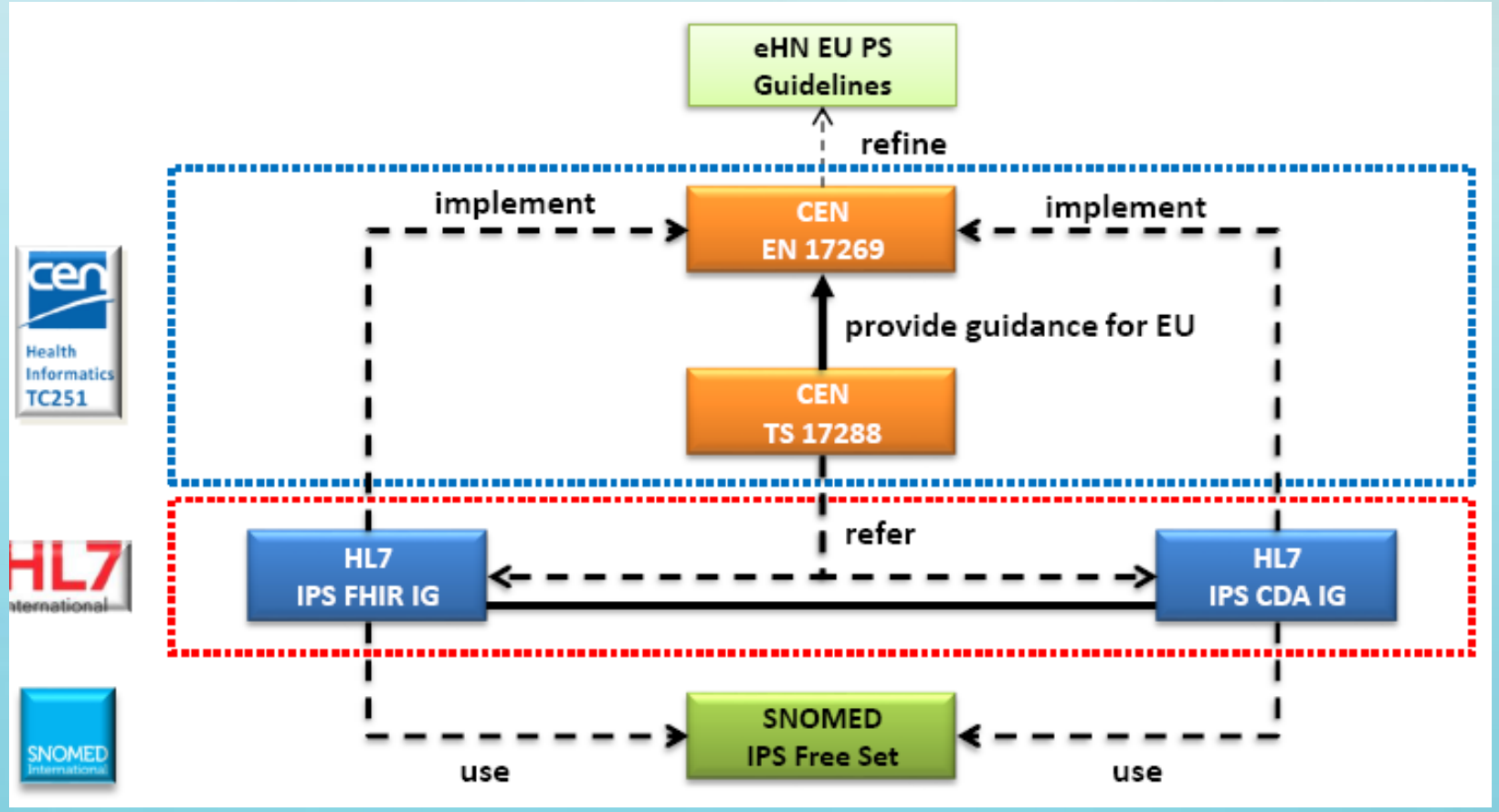
FINAL DRAFT
FprEN 17269

July 2019

TECHNICAL SPECIFICATION
SPÉCIFICATION TECHNIQUE
TECHNISCHE SPEZIFIKATION

FINAL DRAFT
FprCEN/TS 17288

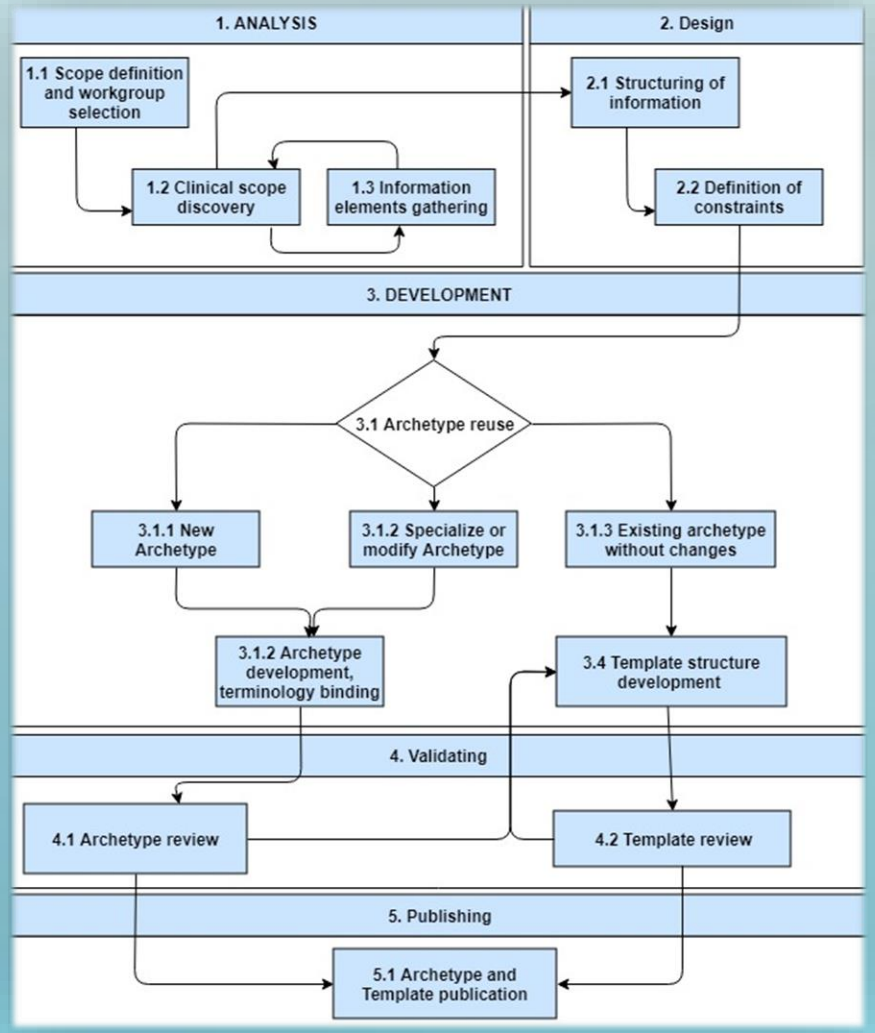
September 2018



2.1 International Patient Summary (IPS) Case study objectives

- Implement the IPS Medication section in prEN 17269 both in terms of CEN 13606 Archetype Object Model (AOM)and in terms of the openEHR AOM.
- Explore the compatibility of the obtained CEN 13606 Archetype conceptual design with respect to the requirements of an openEHR engine for running openEHR Operational templates.
- Explore the W3C XML Schemas of IPS archetype conceptual models with respect to potential practical implementations of the proposed standard.
- Develop a client- server application for testing the openEHR Operational template on an openEHR engine in a local and cloud environment.
- Propose a methodology for transforming a CEN 13606 or openEHR archetype conceptual models into a format that enables the creation of archetype instances compatible with Native XML Databases(NXD).

2.1 International Patient Summary (IPS) Case study methods



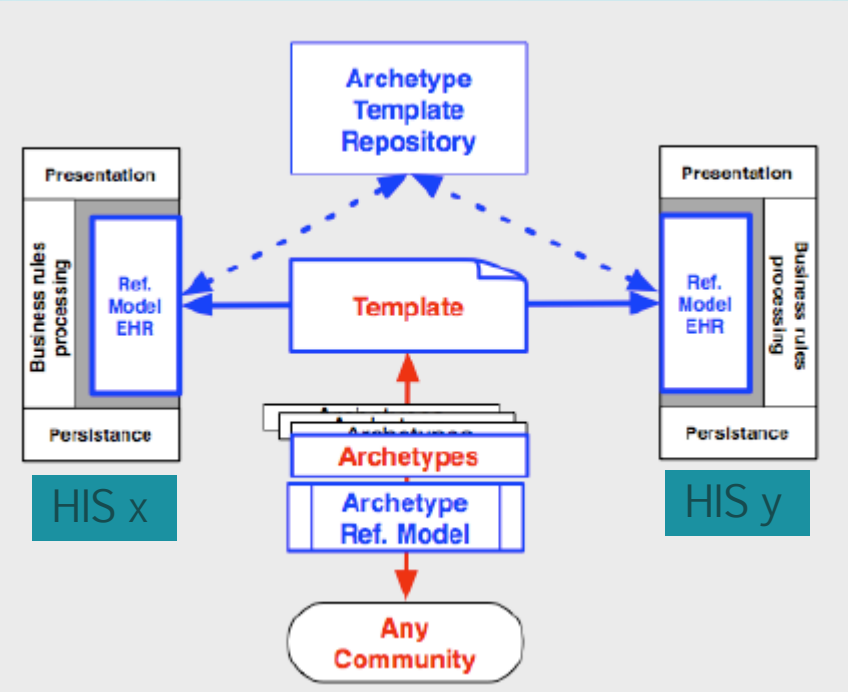
Methodology for Archetype design

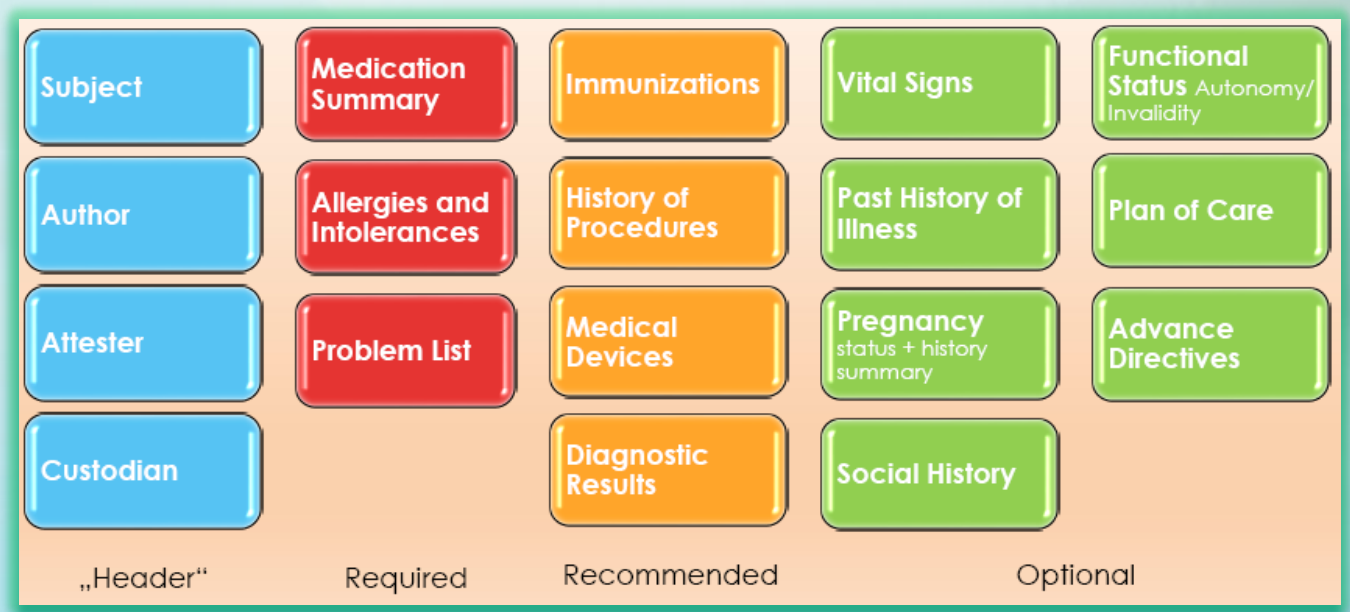
IPS Design with Archetype Concepts

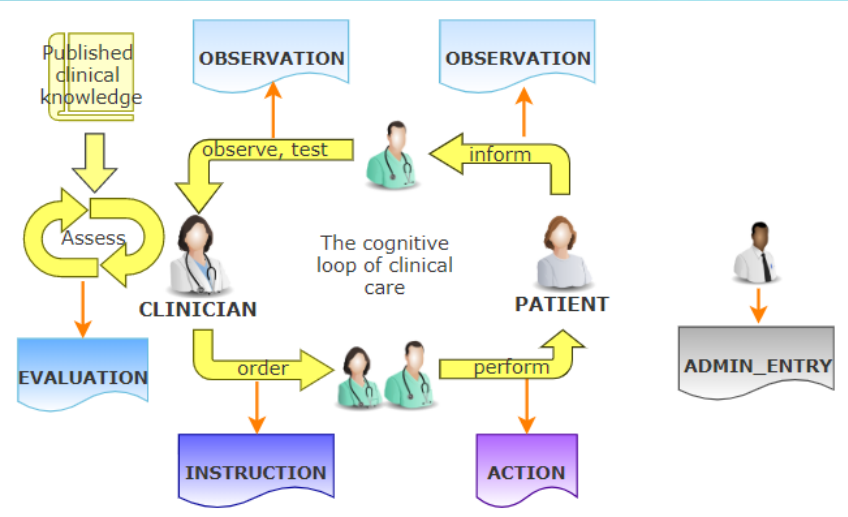
The Archetype Paradigm
(ISO EN 13606:1, ISO EN 13606:2, openEHR)

- ✓ Based on a dual Information model comprising:
  - A Reference Model (RM)- a generic model for all possible kinds of records
  - Archetype Object Model- specification of constraints to the RM
- ✓ Makes 'Plug-and-Play' semantic interoperability possible
- ✓ One general health record specification can be implemented only once

IPS Design with Archetype Concepts



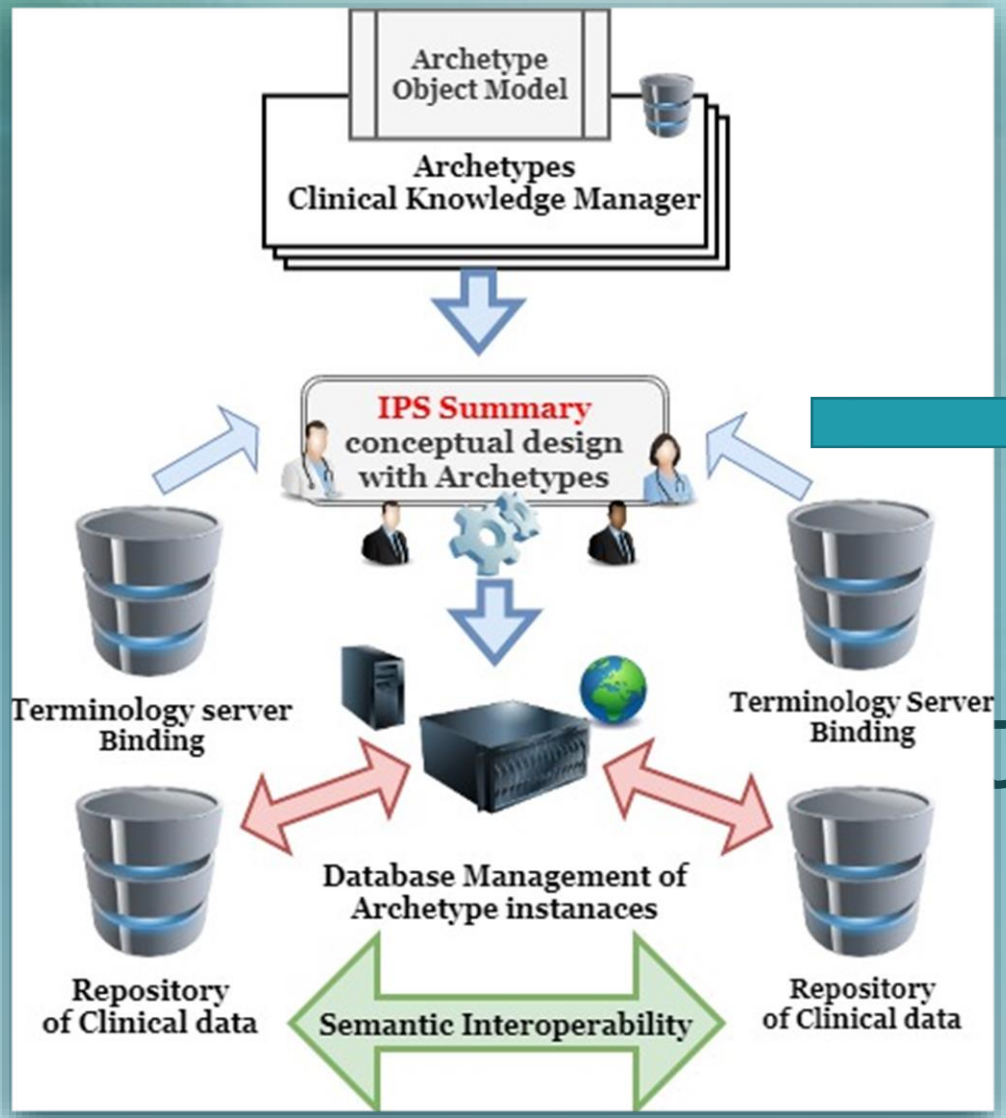




18.1 Overview Description for MEDICATION SUMMARY

Table 26 — Medication Summary Overview

Patient clinical data						
Hierarchy: H1	H2	H3	H4 -- H7	Conformance	Description	Further Details
IPS Section: MEDICATION SUMMARY (PART 1) Synonyms: None Acronyms: None				M	Every PS conformant to IPS must contain this IPS section.	#1
	Medication Summary content status			C	Coded Element	#2
	List of medication			C	List	#3
		Medication		M	Label Concept	#4

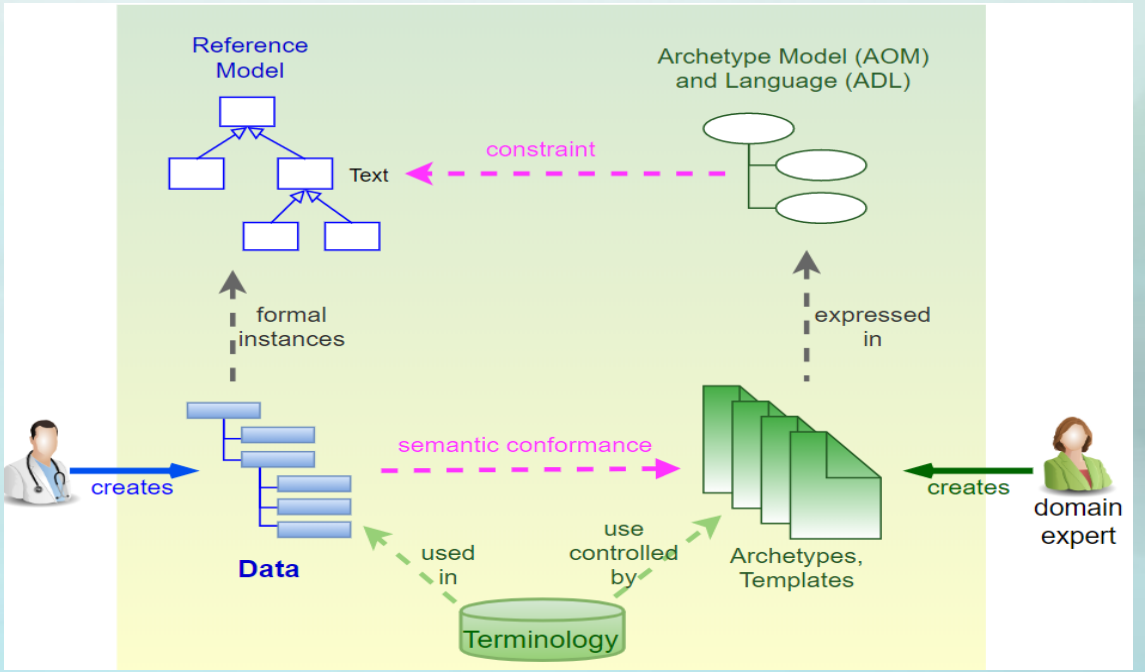


tle

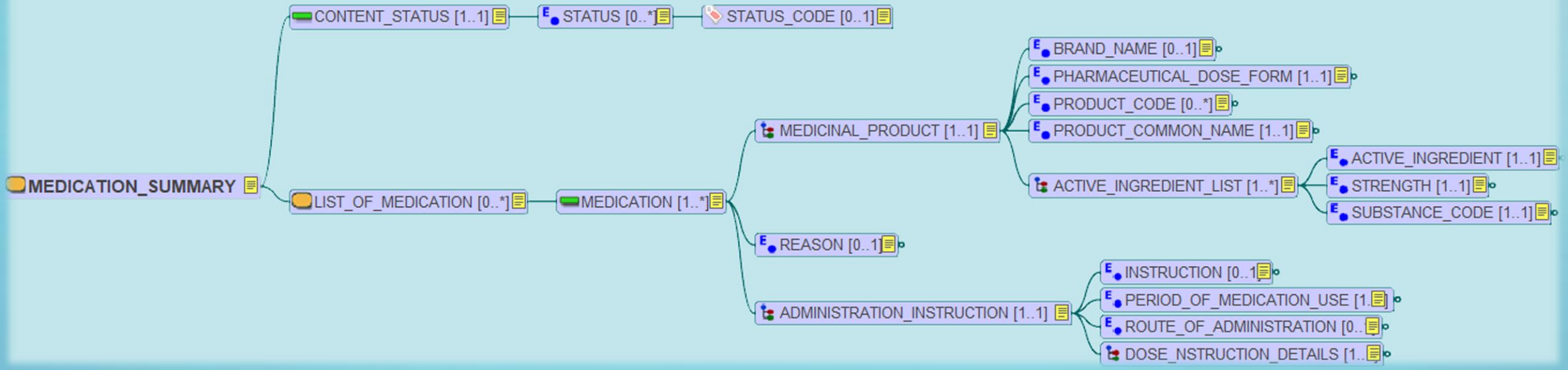
Table 27 — Medication Summary and IDMP

Patient clinical data							
Hierarchy: H3	H4	H5	H6	H7	Conformance	Description	Further Details
IPS Section: MEDICATION SUMMARY (PART 2 beginning with H3 level) Medication					M	Part 2 comprises a paired list of Medicine and Administration. Labelled Concept	#4
	Reason				O	Label Concept	#5
	Medicinal Product				R	Label Concept	#6
		Product Code			O	Coded Element	#7
		Product Common Name (and Strength)			RK	String	#8
		Pharmaceutical dose form			R	Coded Element	#9
		Brand name			O	String	#10
		Active ingredient List			R	List	#11
			Active Ingredient		R	Label Concept	
				Substance code	R	Coded Element	#12
				Strength	R	Ratio	#13
	Administration Instruction				R	Label concept	#14
		Instruction			O	Text	#15
		Period of Medication Use			R	Period	#16
		Route of Administration			O	Coded Concept	
		Dose Instruction			R	Label Concept	
			No. of units per intake		R	Range or Quantity	#17
			Frequency of intake		R	General Time Specification	#18

```
archetype (adl_version=1.4)
    CEN-EN13606-SECTION.MEDICAL_SUMMARY.v1

concept
    [at0000]

language
    original_language = <[ISO_639-1::en]>

description
    original_author = <
        ["date"] = <"2019-06-10">
    >
    lifecycle_state = <"Draft">
    details = <
        ["en"] = <
            language = <[ISO_639-1::en]>
            other_details = <
                ["lastExportDate"] = <"06/10/2019 11:36:07">
            >
        >
    >
    other_details = <
        ["lastExportDate"] = <"06/10/2019 14:25:56">
    >

definition
    SECTION[at0000] occurrences matches {1..1} matches {  -- MEDICATION_SUMMARY
        members existence matches {0..1} cardinality matches {0..*; ordered} matches {
            ENTRY[at0001] occurrences matches {1..1} matches {  -- CONTENT_STATUS
                items existence matches {0..1} cardinality matches {0..*; unordered; unique} matches {
                    ELEMENT[at0002] occurrences matches {0..*} matches {  -- STATUS
                        value existence matches {0..1} matches {
                            CODED_TEXT[at0003] occurrences matches {1..1} matches {  -- STATUS_CODE
                                codedValue existence matches {1..1} matches {*}
                                originalText existence matches {0..1} matches {/.*/}
                                charset existence matches {0..1} matches {
                                    CS[at0015] occurrences matches {0..1} matches {  --
```

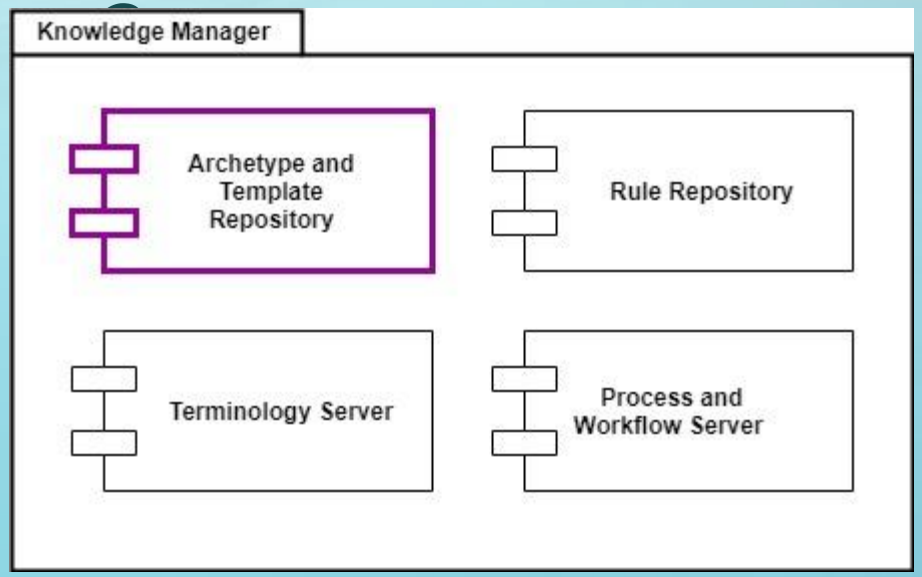


Dual Information Model



2.2 International Patient Summary (IPS). Case study Results

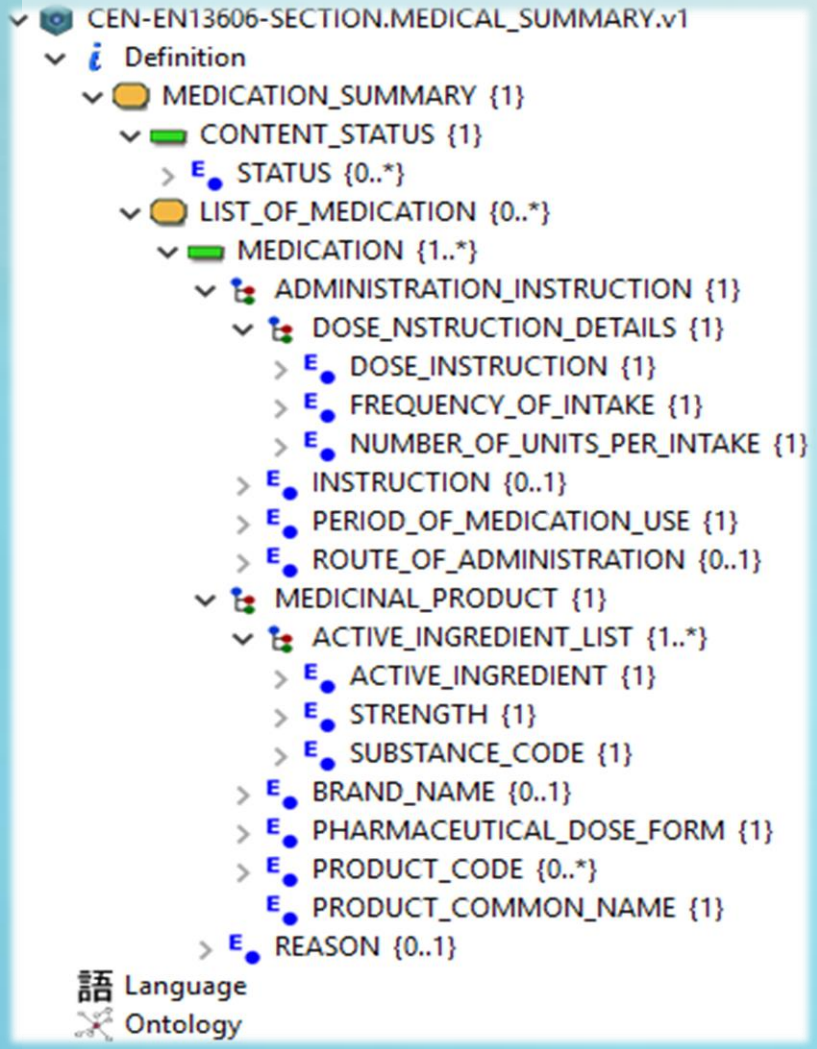
- IPS SECTION: Medication summary
  - INSTRUCTION
    - Medical summary
      - Fillin Medication summary details Δ *NAME (from 'Current Activity')*
        - description
          - CONTENT_STATUS Δ *[0..1] to [1..1],*
          - LIST_OF_MEDICATION
            - MEDICATION
              - T REASON
              - MEDICINAL_PRODUCT Δ *[0..1] to [1..1],*
                - T PRODUCT_CODE
                - T PRODUCT_COMMON_NAME Δ *[0..1] to [1..1],*
                - T PHARMACEUTICAL_DOSE_FORM Δ *[0..1] to [1..1],*
                - T BRAND_NAME
                - ACTIVE_INGREDIENT_LIST Δ *[0..1] to [1..1],*
                  - ACTIVE_INGREDIENTS Δ *[0..1] to [1..1],*
                    - Q STRENGTH Δ *[0..1] to [1..1],*
                    - T SUBSTANCE_CODE Δ *[0..1] to [1..1],*
                    - T ACTIVE_INGREDIENT
              - ADMINISTRATION_INSTRUCTION Δ *[0..1] to [1..1],*
                - T INSTRUCTION
                - PERIOD_OF_MEDICATION_USE Δ *[0..1] to [1..1],*
                - T ROUTE_OF_ADMINISTRATION
                - DOSE_NSTRUCTION_DETAILS Δ *[0..1] to [1..1],*
                - NUMBER_OF_UNITS_PER_INTAKE Δ *[0..1] to [1..1],*
                - FREQUENCY_OF_INTAKE Δ *[0..1] to [1..1],*
                - T DOSE_INSTRUCTION

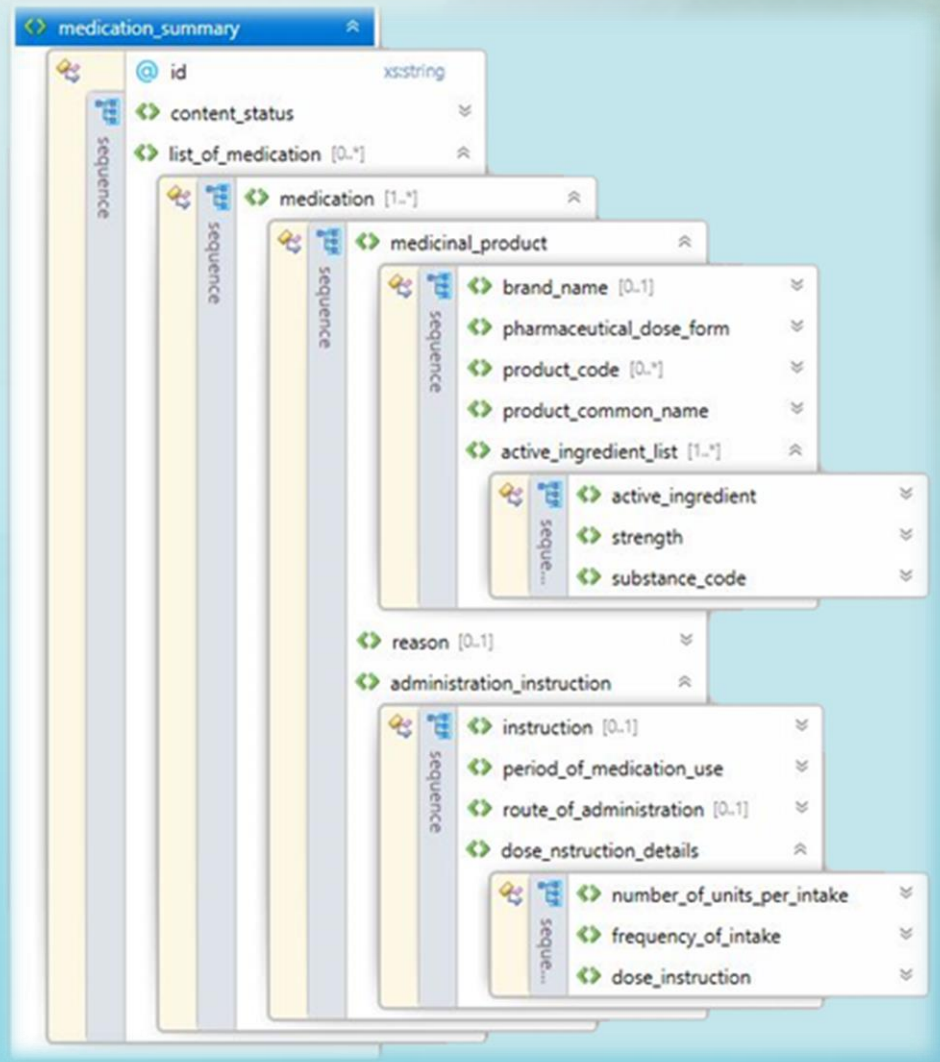


openEHR
Clinical Knowledge Manager

2.2 International Patient Summary (IPS). Case study Results

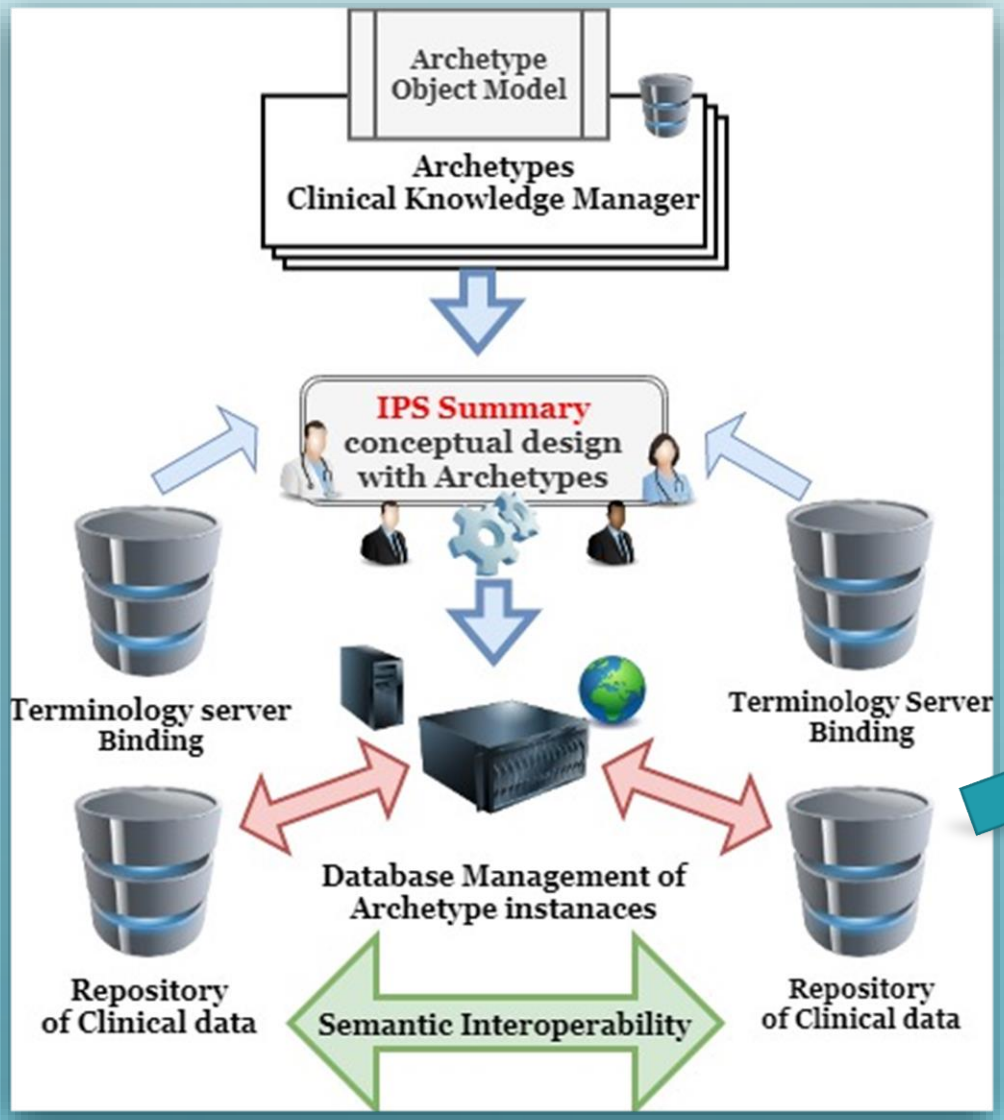
Conceptual Design With Archetypes

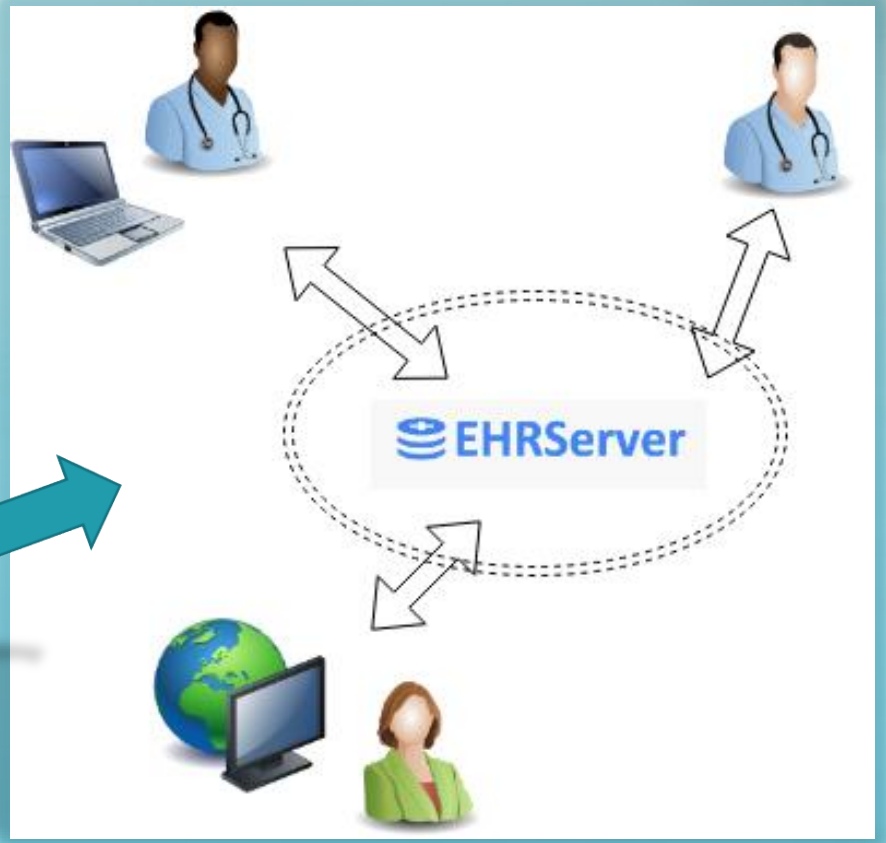




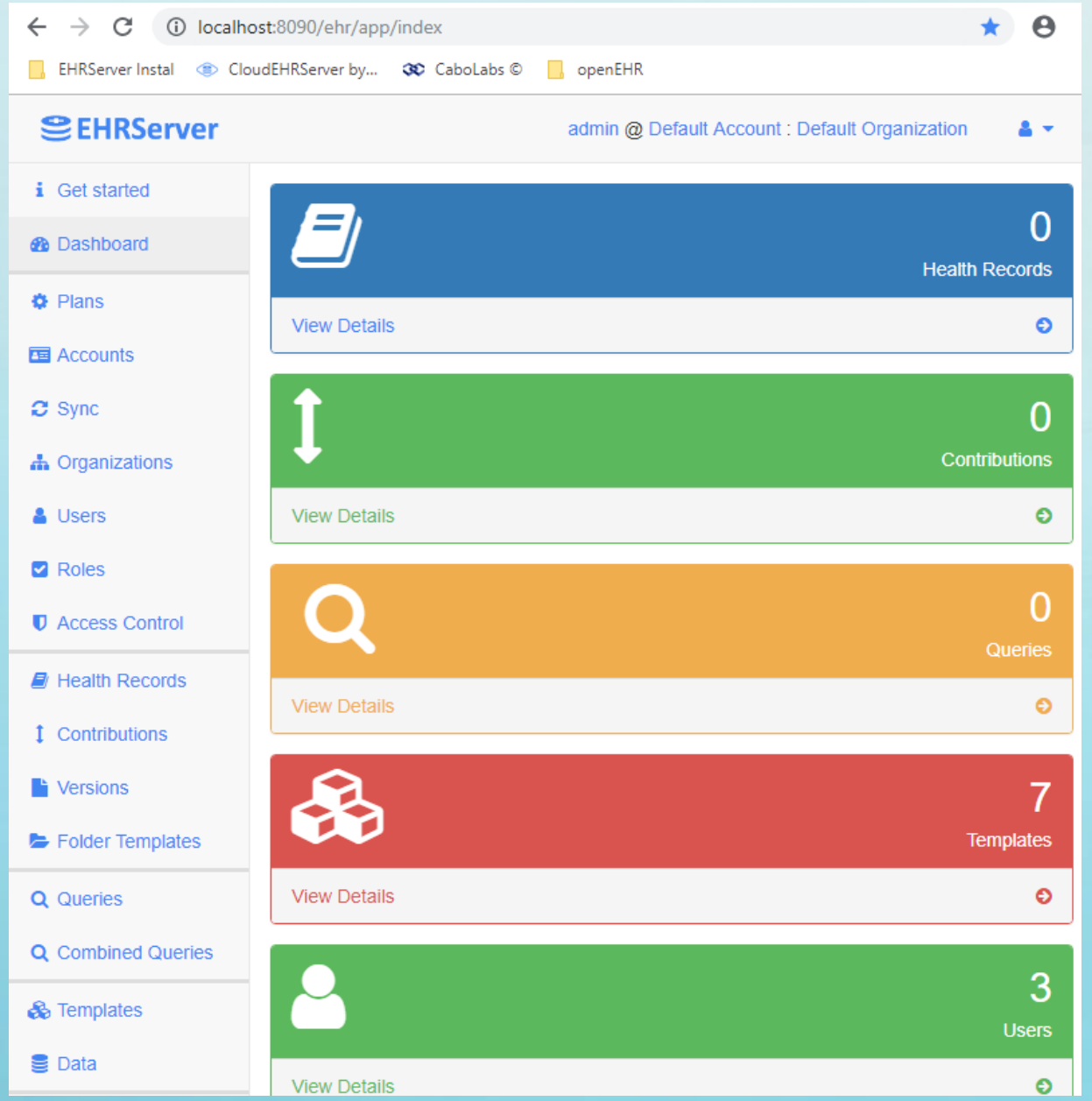
2.2 International Patient Summary (IPS). Case study Results

Software Implementation





2.2 International Patient Summary (IPS). Case study Results

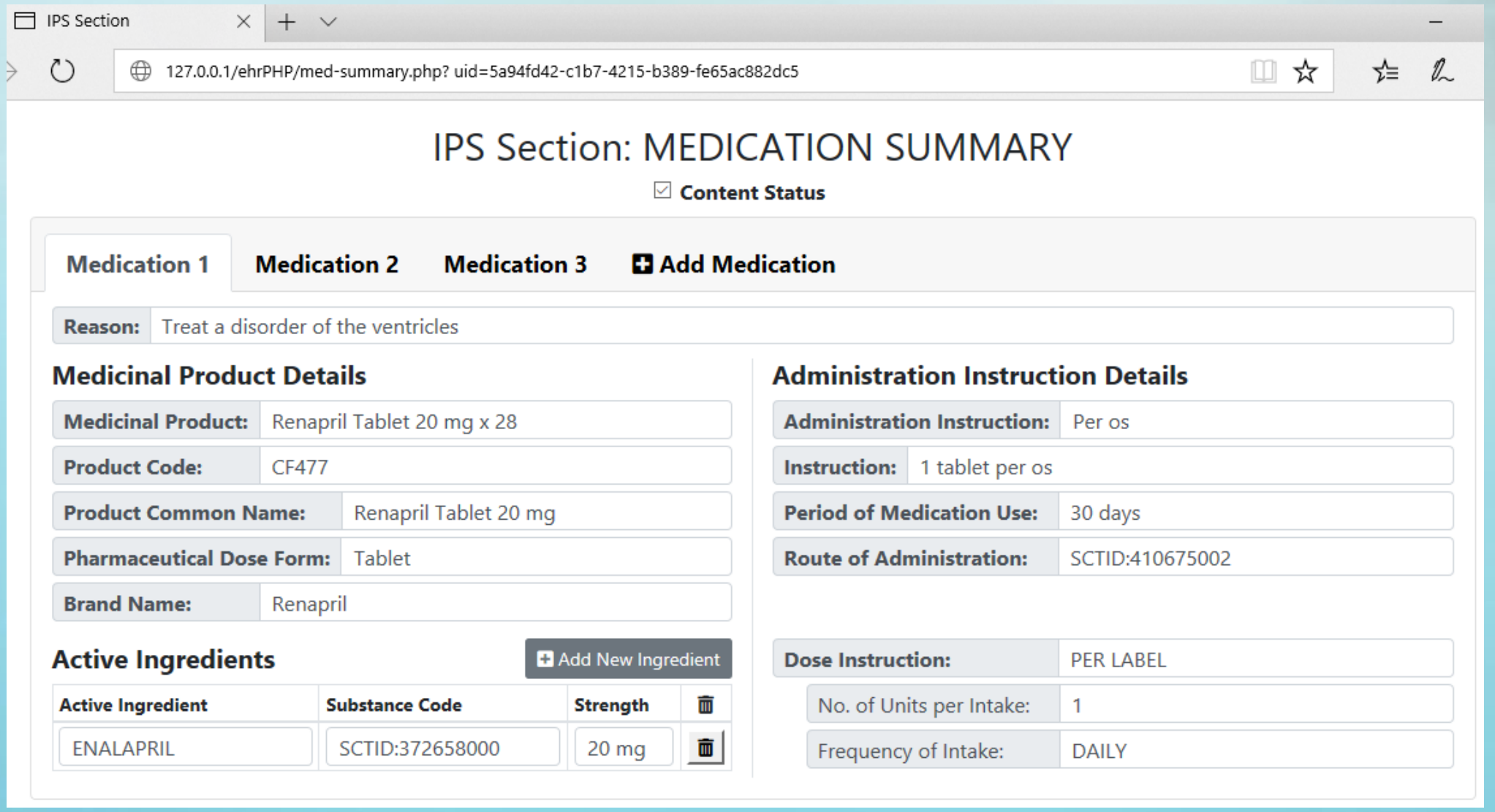


openEHR engine

- ✓ data repository with a REST web services API
- ✓ Document structures defined as Archetypes
- ✓ Archetypes embedded into openEHR Operational Templates
- ✓ Provides an Archetype Query Language
- ✓ Agnostic about what data is stored in its database

2.2 International Patient Summary (IPS). Case study Results

openEHR Frontend Application- prototype



2.3 International Patient Summary (IPS). Case study Summary

Novel results- video presentation in YouTube:

- ✓ Applied the Archetype paradigm in software applications that implement the draft version of the IPS standard.
- ✓ Proposed a methodology for applying the proposed IPS standard in use cases where semantic interoperability is a requirement.
- ✓ Provide public access to IPS implementation in archetypes in ISO 13606 and openEHR.
- ✓ Established difficulties in implementing semantic interoperability with IPS due to the lack of free access to terminology servers
- ✓ The obtained archetype conceptual model is ready for reuse in software applications
- ✓ Established that it is not possible to convert ISO 13606 archetypes directly into openEHR archetypes.
- ✓ Relational databases are more suitable for openEHR implementations of the IPS, while ISO 13606 implementations of the IPS are more productive with Native XML databases

2.3 International Patient Summary (IPS). Case study Summary

- ✓ The current draft versions of EN 17269 and CEN/TS 17288. **make use exclusively** of HL7, FHIR and HL7 CDA technologies
- ✓ The analysis of these documents shows that they are built with **a focus on top of the Message paradigm** and provide examples where semantic interoperability is not taken in consideration.
- ✓ **The obtained results provide evidence that EN 17269 and CEN/TS 17288 can be enabled to provide semantic interoperability**, where the IPS design **is based on the Archetype paradigm**.
- ✓ **The obtained results are novel** because the Archetype paradigm is not considered in draft version of the IPS standard EN 17269 and in the existing literature. They serve to **extend the practical experience in cross-border sharing of clinical data** represented **in terms of semantic interoperability of archetype concepts**

Video presentation in YouTube

Kovachev, P., 2020. *openEHR implementation of International Patient Summary based on the CEN 17269:2019 standard*. [Online]

2.3 International Patient Summary (IPS). Case study Summary

Tcharaktchiev D, Krastev E, Petrossians P, Abanos S, Kyurkchiev H, Kovatchev P. "Cross-Border Exchange of Clinical Data Using Archetype Concepts Compatible with the International Patient Summary". Stud Health Technol Inform. IOS Press 2020 Jun 16; 270:552-556., doi: 10.3233/SHTI200221. (open access)

Evgeniy Krastev, Petko Kovatchev, Dimitar Tcharaktchiev, Simeon Abanos, "Primary Use Case Implementation of International Patient Summary on openEHR Platform", Proc. of 12th International Conference e-Health 2020 (EH 2020), Computer Science and Information Systems Series, pp. 167- 174, IADIS Press, ISBN: 978-989-8704-18-4, (open access)

Evgeniy Krastev, Dimitar Tcharaktchiev, Lyubomir Kirov, Petko Kovatchev, Simeon Abanos, and Alexandrina Lambova, "Software Implementation of the EU Patient Summary with Archetype Concepts", In Proceedings of GLOBAL HEALTH 2019, The Eighth International Conference on Global Health Challenges, Porto, Portugal, from September 22, 2019 to September 26, 2019 pp. 8- 13 ISSN: 2308-4553, ISBN: 978-1-61208-742-9

Evgeniy Krastev, Dimitar Tcharaktchiev, Petko Kovatchev, Simeon Abanos, "International Patient Summary Standard Based on Archetype Concepts" International Journal on Advances in Life Sciences, ISSN 1942-2660 vol. 12, no. 1 & 2, year 2020, 34 :46, (open access)

3. OMOP Common Data Model

European Health Data & Evidence Network (EHDEN):

Objective:

- ✓ Create an **EU-wide ecosystem** for federated analyses of observational data **by mapping** to the **OMOP CDM** an excess of 100+ millions of anonymous health records
- ❑ Supported by a large number of certified **Small-to-Medium-sized Enterprises (SME)**, **OHDSI Software and Tools**, **OMOP Standardized vocabularies** and **Technical infrastructure Services**.
- ❑ EHDEN Data Partner
  National Scientific Program E-Health in Bulgaria

3. OMOP Common Data Model

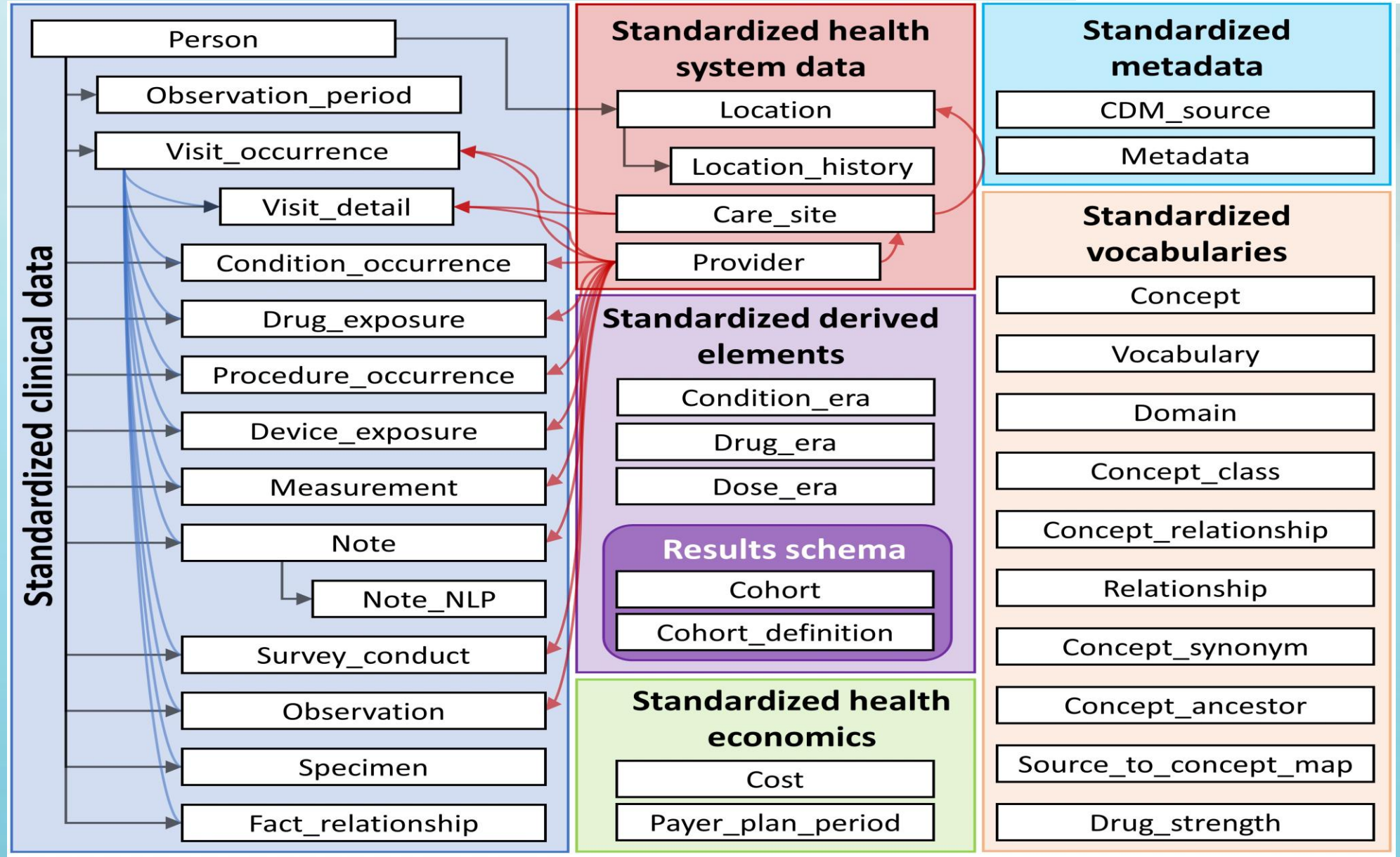
Digital health technologies produce huge amounts of data related to patient health collected as part the execution of routine healthcare services under real-world conditions.

Data collected from such sources is collectively known as **observational data** (OD).

The **Observational Medical Outcomes Partnership** (OMOP) **Common Data Model** (CDM) is an open community data standard, designed to **standardize the structure and content of observational data** and to **enable efficient analyses** that can **produce reliable evidence**.

It is a **unified database model** that allows to integrate various OD sources including EHRs based on this standard.

3. OMOP Common Data Model



3. OMOP Common Data Model

The Common Data Model (CDM) **defines the structure and content of data** to provide **significant and reproducible results.**

CDM Design Principles:

- Suitability for purpose
- Data protection:
- Design of domains
- Rationale for domains
- Standardized Vocabularies
- Reuse of existing vocabularies
- Maintaining source codes
- Technology neutrality
- Scalability
- Backwards compatibility

3. OMOP Common Data Model

- **Suitability for purpose**: The CDM aims to provide data organized in a way optimal for analysis, rather than for the purpose of addressing the operational needs of health care providers or payers.
- **Data protection**: All data that might jeopardize the identity and protection of patients, such as names, precise birthdays etc. are limited. Exceptions are possible when the research expressly requires more detailed information, such as precise birth dates for the study of infants.
- **Design of domains**: The domains are modeled in a person-centric relational data model, where for each record the identity of the person and a date is captured as a minimum. Here, a relational data model is one where the data is represented as a collection of tables linked by primary and foreign keys.
- **Rationale for domains**: Domains are identified and separately defined in an entity-relationship model if they have an analysis use case (conditions, for example) and the domain has specific attributes that are not otherwise applicable. All other data can be preserved as an observation in the observation table in an entity-attribute-value structure.

3. OMOP Common Data Model

- **Standardized Vocabularies**: To standardize the content of those records, the CDM relies on the Standardized Vocabularies containing all necessary and appropriate corresponding standard healthcare concepts.
- **Reuse of existing vocabularies**: If possible, these concepts are leveraged from national or industry standardization or vocabulary definition organizations or initiatives, such as the National Library of Medicine, the Department of Veterans' Affairs, the Center of Disease Control and Prevention, etc.
- **Maintaining source codes**: Even though all codes are mapped to the Standardized Vocabularies, the model also stores the original source code to ensure no information is lost.
- **Technology neutrality**: The CDM does not require a specific technology. It can be realized in any relational database, such as Oracle, SQL Server etc., or as SAS analytical datasets.
- **Scalability**: The CDM is optimized for data processing and computational analysis to accommodate data sources that vary in size, including databases with up to hundreds of millions of persons and billions of clinical observations.

3. OMOP Common Data Model

Backwards compatibility:
All changes from previous CDMs are clearly delineated in the github repository

- ✓ The CDM is considered a "**person-centric**" model, meaning that all clinical Event tables are linked to the PERSON table. Together with the date or start date this allows for a longitudinal view on all healthcare-relevant Events by person. The exceptions from this rule are the standardized health system data tables, which are linked directly to Events of the various domains.
- ✓ Schemas, or database users in some systems, **allow for separation between read-only and read-write tables**. The clinical Event and vocabulary tables are in the "CDM" schema and are considered read-only to the end user or analytic tool. Tables that need to be manipulated by web-based tools or end users are stored in the "Results" schema.
- ✓ The **CDM is platform-independent**. Data types are defined generically using ANSI SQL data types (VARCHAR, INTEGER, FLOAT, DATE, DATETIME, CLOB). Precision is provided only for VARCHAR. The CDM does not prescribe the date and datetime format.
- ✓ **CDM provides standardized DDL SQL for creating schemas** according CDM model for most RDBMs like PostgreSQL, Oracle, MSSQL, BigQuery, Amazon Redshift and other.

3. OMOP Common Data Model

CDM General conventions:

- General Conventions of the Model
- General Conventions of Schemas
- General Conventions of Data Tables

3.1 OMOP Common Data Model Methods

A. Preparing the CDM database

B. Creating the required database users;

C. Creating the OMOP CDM tables with the Common Data Model/MSSQL DDL scripts;

D. Importing the standard OMOP standarizied vocabularies from http://athena.ohdsi.org

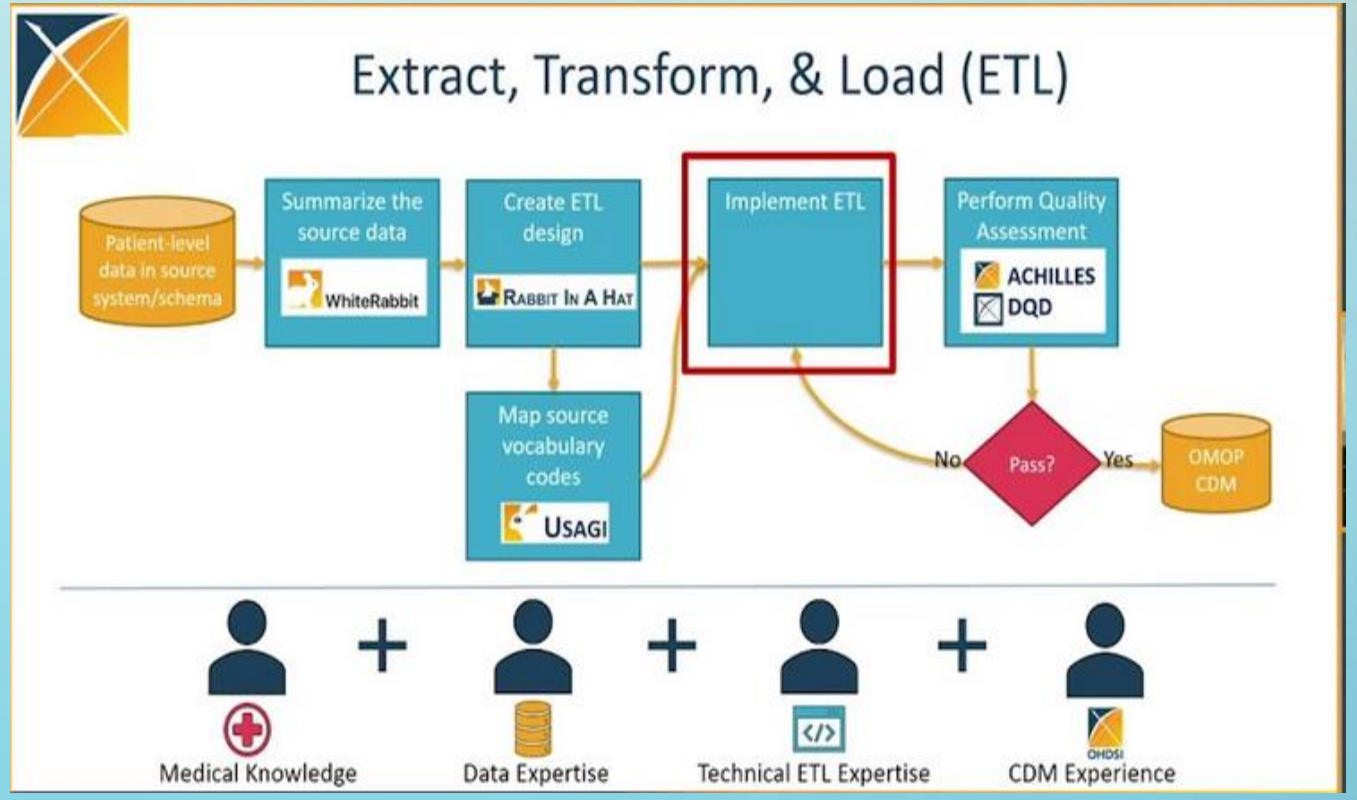
- All the OHDSI sources are available at http://github.org/OHDSI.

3.1 OMOP Common Data Model Methods

ETL - Extract Transform Load

In order to get from the native/raw data to the OMOP Common Data Model (CDM) we have to create an extract, transform, and load (ETL) process with these **steps**:

1. Design the ETL
2. Create the Code Mappings
3. Implement the ETL
4. Quality Control



3.1 OMOP Common Data Model Methods

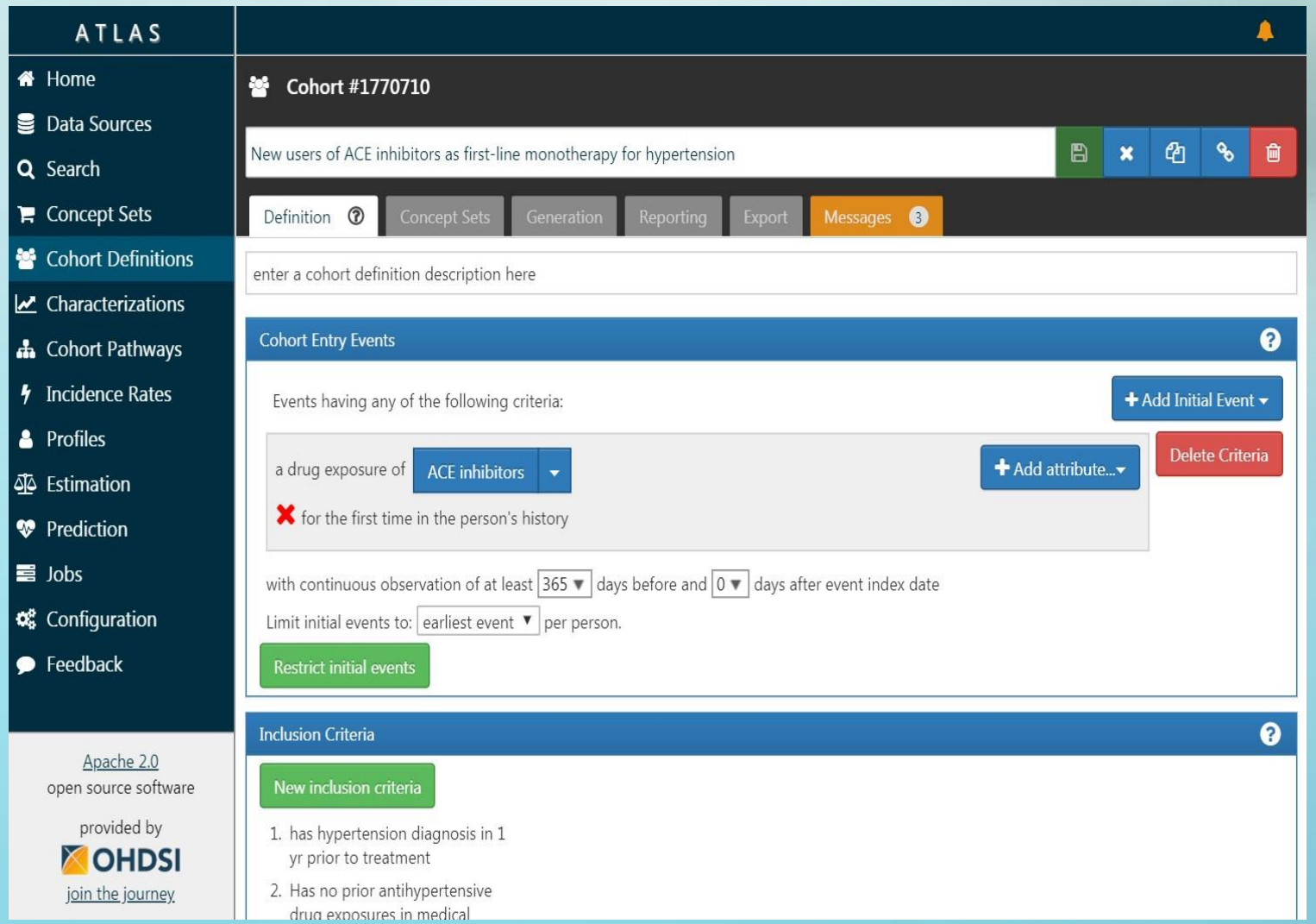
ETL - Extract Transform Load

1. To **initiate an ETL process** on a database we need to understand source data, including the tables, fields, and content. The White Rabbit software from OHDSI to perform a scan of the source data. The scan generates a report used as a reference when designing the ETL. With the White Rabbit scan in hand, we have a clear picture of the source data. We also know the full specification of the CDM.
2. With Usagi **from OHDSI tools** we perform manual process of creating a code mappings with standard source codes to Vocabulary concepts
3. Once the design and code mappings are completed, the ETL process is implemented with ETL-CDM Builder. As result, we have **CDM relevant database** populated with the **data from the source database**

For the extract, transform, load process, **quality control is iterative**.
The typical pattern is to **Write logic** > **Implement logic** > **Test** logic > **Fix**

3.2 OMOP Common Data Model Results

The most convenient and precise approach to perform observational study against CDM database is to **use** [ATLAS] free, publicly available, web-based tool developed by OHDSI that facilitates the design and execution of analyses on standardized, patient level, observational data in the CDM form.



3.2 OMOP Common Data Model Results

ATLAS is deployed as a web application in combination with the OHDSI Web API and is hosted on Apache Tomcat and could be deployed and started as Docker container or cloud service.

- **Data Sources** - provides the capability review descriptive, standardized reporting for each of the configured data sources.
- **Vocabulary Search** - provides the ability to search and explore the OMOP standardized vocabulary.
- **Concept Sets** provides the ability to create collections of logical expressions that can be used to identify a set of concepts to be used throughout your standardized analyses.
- **Cohort Definitions** - ability to construct a set of persons who satisfy one or more criteria for a duration of time.

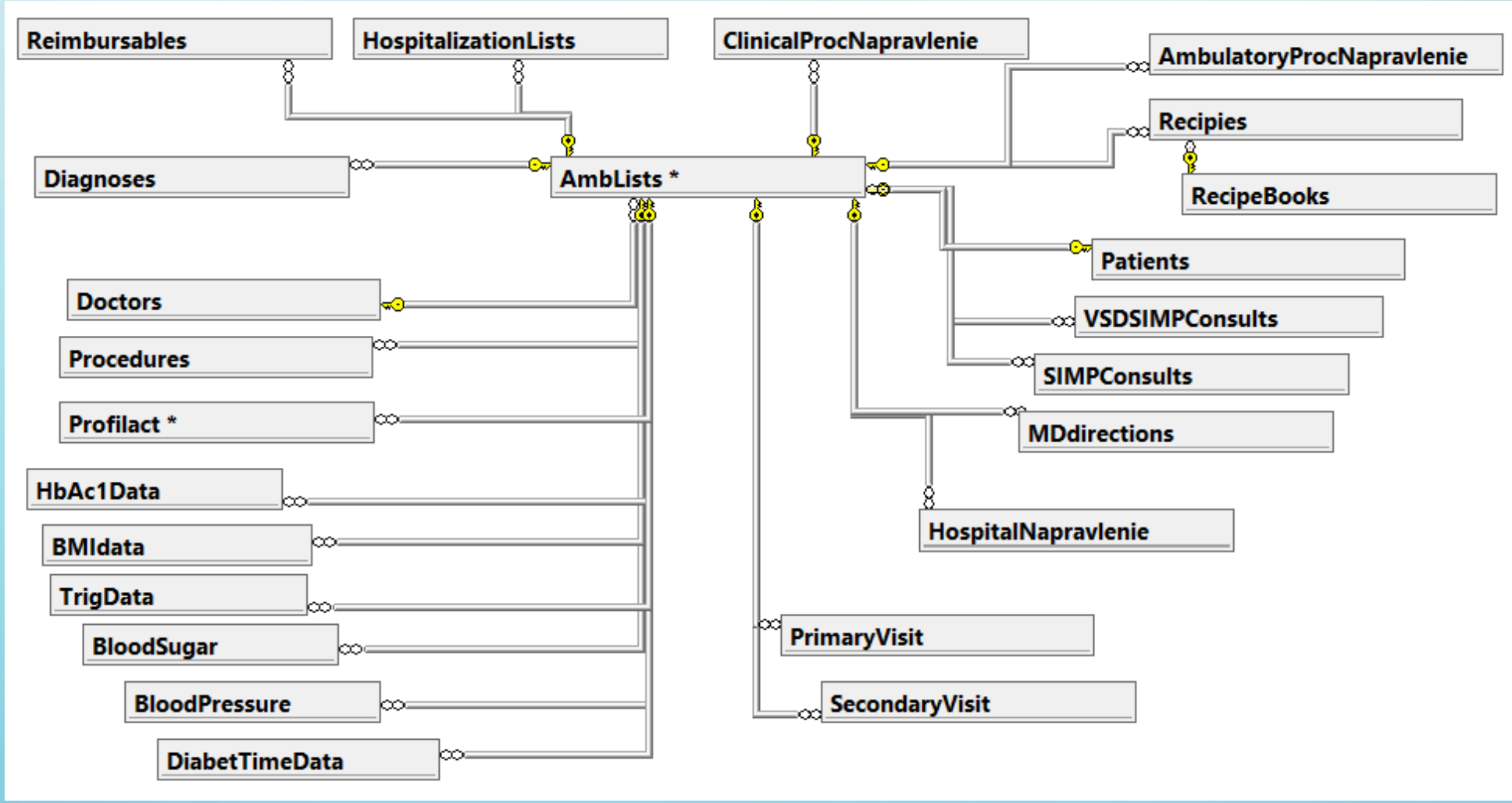
3.2 OMOP Common Data Model Results

- **Characterizations-** an analytic capability that allows you to look at one or more cohorts and to summarize characteristics about those patient populations.
- **Cohort Pathways** Cohort pathways is an analytic tool that allows you to look at the sequence of clinical events that occur within one or more populations.
- **Incidence Rates-** a tool that allows you to estimate the incidence of outcomes within target populations of interest.
- **Profiles-** tool that allows exploring of an individual patients longitudinal observational data to summarize what is going on within a given individual.
- **Population Level Estimation-** a capability that allows to define a population level effect estimation study using a comparative cohort design whereby comparisons between one or more target and comparator cohorts can be explored for a series of outcomes.

3.2 OMOP Common Data Model Results

- **Patient Level Prediction-** allows to apply machine learning algorithms to conduct patient level prediction analyses whereby you can predict an outcome within any given target exposures.
- **Jobs** - used to explore the state of processes that are running through the WebAPI.
- **Configuration** - to review the configured data sources that have been in the source configuration section.

Entity relationship diagram of the Source database taken from the Bulgaria Diabetes Register (DIAB2018)



Entity relationship diagram of the Source database taken from the Bulgarian Diabetes Register (DIAB2018)

Source Table	English Name	Description
AmbLists	AmbLists	Outpatient records generated during a visits to GP or Specialist of Ambulatory Care, coded in ICD9.
Diagnoses	Diagnoses	Diagnoses set in Outpatient record
Doctors	Doctors	Doctors data
HospitalizationLists	Hospitalization Lists	List with hospitalization directions related to an outpatient record
HospitalNapravlenie	HospitalNapravlenie	Direction for hospitalization
MDdirections	MDdirections	Direction for medical examination by Specialist in Ambulatory care
Patients	Patients	Patient description data
PrimaryVisit	PrimaryVisit	Initial visit to a Specialist in Ambulatory Care
Procedures	Procedures	Procedures assigned in Outpatient record
Profilact	Profilact	Describes a visit for Disease prevention
RecipeBooks	RecipeBooks	Recipe Books of Patient
Recipies	Recipies	Recipes for reimbursable medicinal products. These are stored in the patient's Recipe Book
Reimbursables	Reimbursables	Reimbursable medicinal products
SecondaryVisit	SecondaryVisit	Secondary visit to a Specialist in Ambulatory Care
SIMPConsults	SIMPConsults	Specialized Medical care
VSDSIMPConsults	VSDSIMPConsults	Highly Specialized Medical care
BloodPressure	BloodPressure	Blood pressure values measured in mmHg extracted from natural language text description of patient status and examination data
BloodSugar	BloodSugar	Blood sugar profile first value measured in mmol/l extracted from natural text description of patient status and examination data
BMIdata	BMIdata	Body Mass Index data extracted from natural language text description of patient status and examination data. Table includes also Height and Weight measurements in sm and kg, when Height and Weight data are found in the text
HbAc1Data	HbAc1Data	Contains values of HbAc1 extracted from natural language text description of patient status and examination data measured in mmol/mol
DiabetTimeData	DiabetTimeData	Contains the number of years before the illness diabetes has been established for the first time, extracted from natural language text description of patient status and examination data
TrigData	TrigData	Contains values of triglycerides extracted from natural language text description of patient status and examination data and measured in mmol/L

3.2 OMOP Common Data Model Results

In Bulgaria outpatient records are produced by the General Practitioners (GPs) and the Specialists from Ambulatory Care for every contact with the patient.

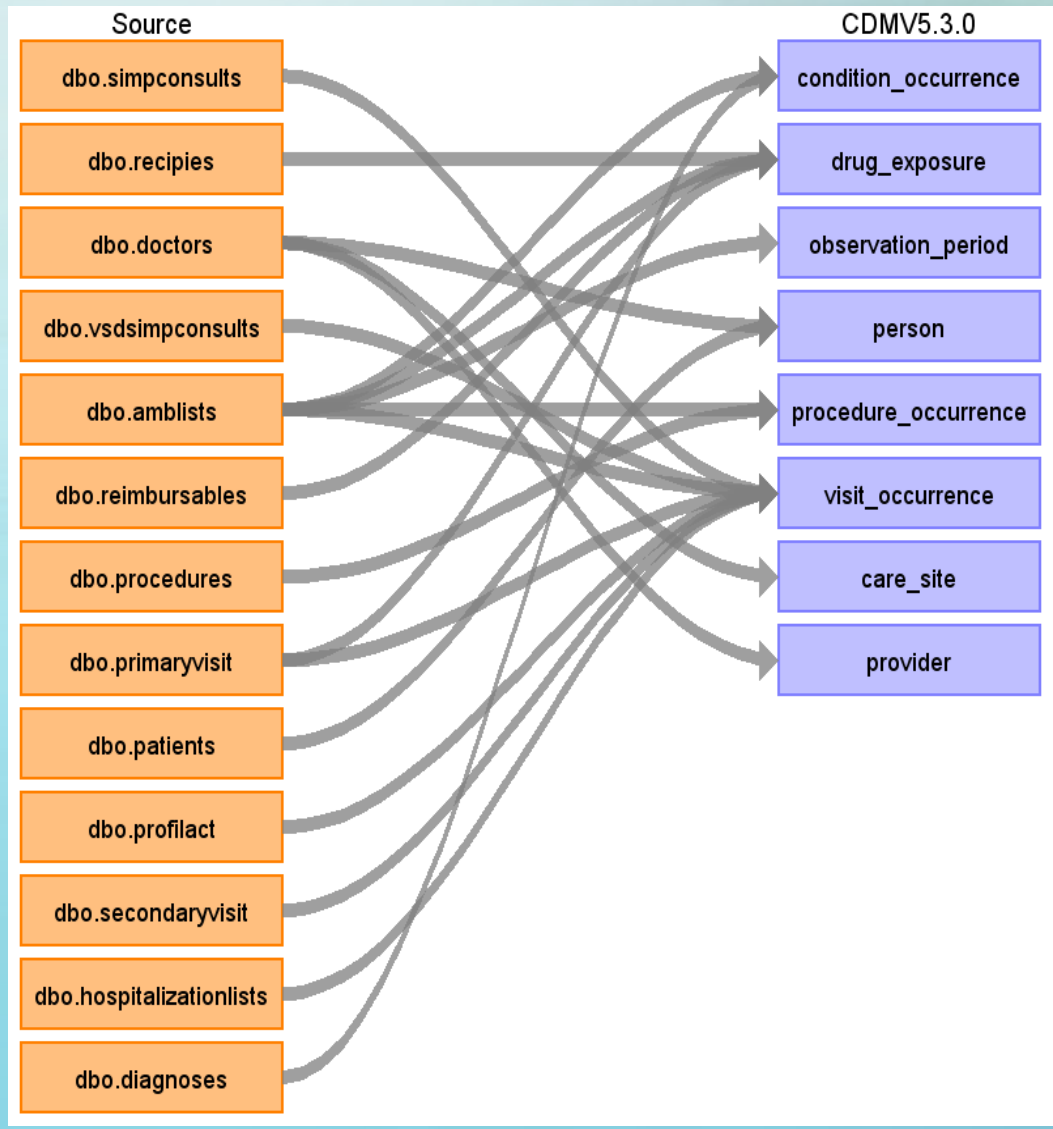
- Outpatient records from patients with diabetes are maintained by the Bulgarian Diabetes Register. These records represent a true example of OD that can be produce valuable evidence for improving the treatment and management the healthcare services for such patients. The outpatient records are semi-structured files with predefined XML schema.
- The source XML documents included more than 1 600 000 pseudo anonymized outpatient records. The most important indicators in the records like Age, Gender, Location, Diagnoses are stored in explicit tags.
- The Case history is presented as free text in the Anamnesis. Additionally, these records include in native text information about the Patient status described the patient state, symptoms, syndromes, patients' height and weight, body mass index (BMI), blood pressure and other clinical concepts.
- The values of clinical tests and lab data are enumerated also as free text in a separate section of the XML document. A special section is dedicated to the prescribed treatment.

3.2 OMOP Common Data Model Results

Source data from DIAB2018 to CDM v. 5.3.0 database mapping using White Rabbit

Database features:

Contains **pseudo anonymized records of 503,753 patients with diabetes** created in ambulatory and specialized care units during 2019- 2019.



3.3 OMOP Common Data Model Summary

- **Common data models** (CDMs) offer a standardized approach for data persistence and exchange. This is especially useful when nowadays clinical data is distributed among heterogeneous sharing systems. Besides, OHDSI provides software tools in support on each stage of the ETL and ensure quality control. Therefore, data CDM possesses all the features of a reliable source for a broad range of statistical analyses.

- This case study presents **initial results of a research work** done with the objective to transfer outpatient records from the Bulgarian Diabetes register into the OMOP CDM. One of the major challenges has been the extraction of clinical data from native text as well as the use of international OMOP concepts to annotate data recorded in a Bulgarian context.

3.3 OMOP Common Data Model Summary

Publications:

Petko Kovachev, Evgeniy Krastev, Dimitar Tcharaktchiev, Emanuil Markov,Ivan Evg. Ivanov, **Conversion of Bulgarian Observational Data to OMOP Common Data Model: InitialResults**. Proceedings of the Information Systems & Grid Technologies: Fifteenth International Conference ISGT'2022, May 27-28, 2022, Sofia, Bulgaria, CEUR Workshop Proceedings, pp. 113-125 (2022)

Krastev E, Tcharaktchiev D, Kovachev P, Abanos S. **Diabetes and Obesity in Bulgaria. Study of a Large Number of Outpatient Records from 2018**. Stud Health Technol Inform. 2022 Jun 29;295:298-301. doi: 10.3233/SHTI220721. PMID: 35773867

Acknowledgment

This research is supported by
the **National Scientific Program eHealth** in Bulgaria

Professor Evgeniy Krastev, PhD
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
eck@fmi.uni-sofia.bg
https://ehealth.fmi.uni-sofia.bg

Professor Dimitar Tcharaktchiev , PhD, MD
Department of Medical Informatics,Medical University-Sofia,
University Hospital of Endocrinology,
2 Zdrave street, Sofia 1431, Bulgaria;
dimitardt@gmail.com
https://ehealth.fmi.uni-sofia.bg

Thanks for your attention!

Questions?

Comments?