Medical Databases
MSc VEMIVIM554E

Medical Data Modeling

István Vassányi, 2017

Outline

1 Introduction

2 Shortcomings of traditional healthcare data models

3 Two simple modeling techniques

4 The two-level concept

5 MSZ 22800 Reference Model overview

6 Application in the national healthcare SOA

7 Summary

Introduction

- The domain model is the cornerstone of an information system
- Standard domain modeling methods
- Most widely used method is UML (class diagrams)
- Typical OLTP system is based on a relational database
- Relational data models can be generated automatically or designed manually

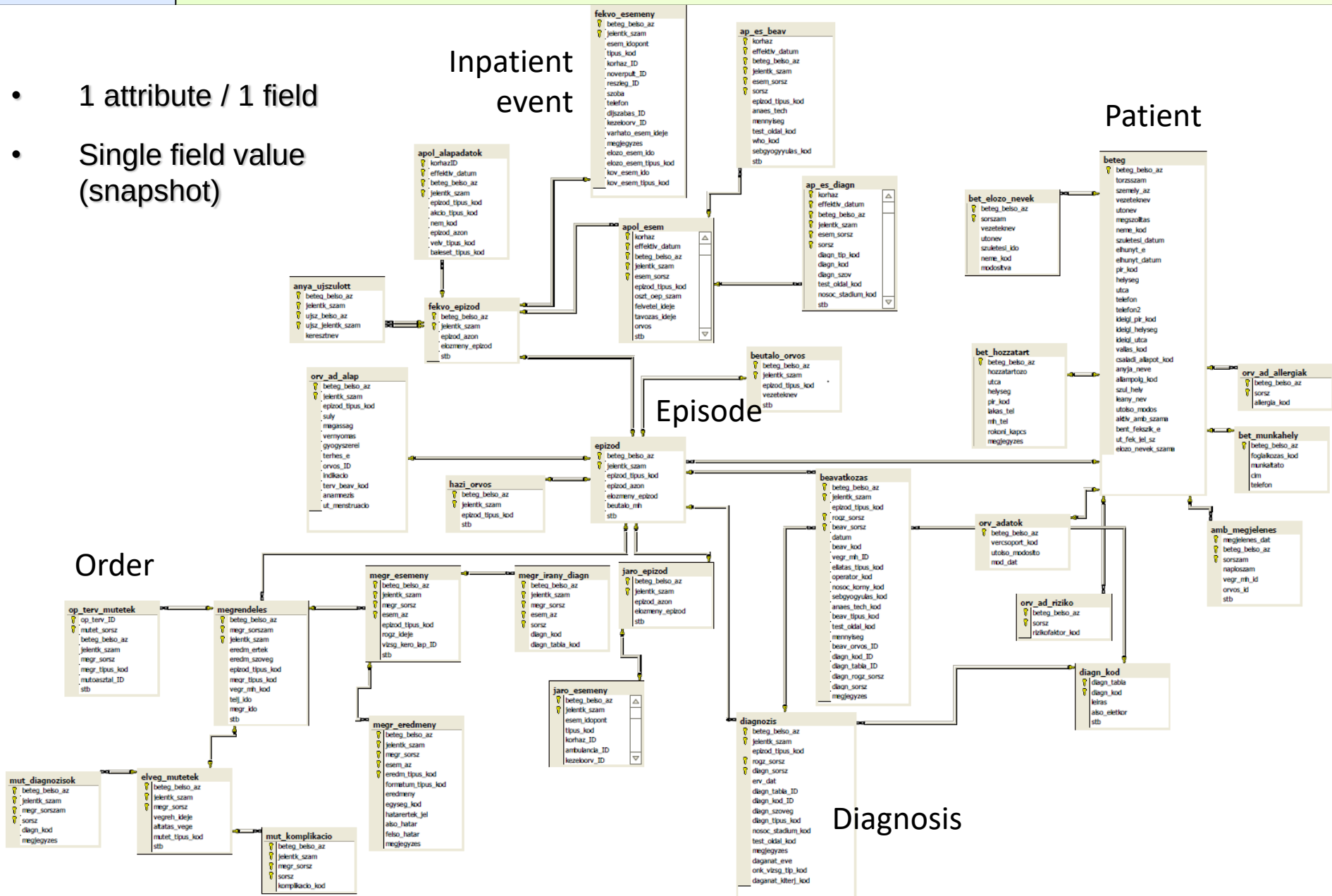
2 Special requirements for healthcare domain models

- Legal and ethical issues
- Data security and data protection
- Reliability and availability, failover protection
- Interoperability, messaging
- Heterogeneous hardware/software environment
- *Medical knowledge and protocols are far more **complex** than other domains → Exceptionally wide semantic gap between the medical professional and the IT professional*
- *Medical knowledge and protocols **change fast** → applications are obsoleted in short time*

Today we will focus only on the *last two issues*.

An example “traditional” HIS data model

- 1 attribute / 1 field
- Single field value (snapshot)



Problems with the “traditional” HIS data model

14. számú melléklet a 43/1999. (III. 3.) Korm. rendelethez [631]

ADATLAP

kórházi (osztályos) ápolási esetről

1. A kórház neve:	**2. Osztály neve:**
3. Osztály azonosítója: / *3/A. Szakmakód:*	*4. Térítési kategória:* / *4/A. Részleges térítés:*
4/B. Részleges térítési díj összege:	**5. Személyazonosító típusa:**
6. Személyazonosító jel:	**7. Törzsszám:**
8. A beteg neve:	**9. Érvényes biztosítás országa vagy a beteg állampolgársága:**
10. Anyja neve:	**11. A beteg születési neve:**
12. A beteg neme:	**13. Születés dátuma:**
14. A kísért beteg törzsszáma:	**14/A. A beteg kíséretének oka:**
16. Lakcím:	**16/A. Lakcím típusa:**
17/A. Beutaló orvos munkahelyének azonosítója	**17/B. Beutaló orvos kódja:**
17/C. Beutalást megalapozó ellátást igazoló adat	**18. A felvétel jellege:** / **19. A felvétel típusa:**
20. A felvétel időpontja:	**22. A távozás időpontja:**
23. A beteg további sorsa: / **24. Az ellátó orvos kódja:**	**25. E-adatlap kitöltés:**

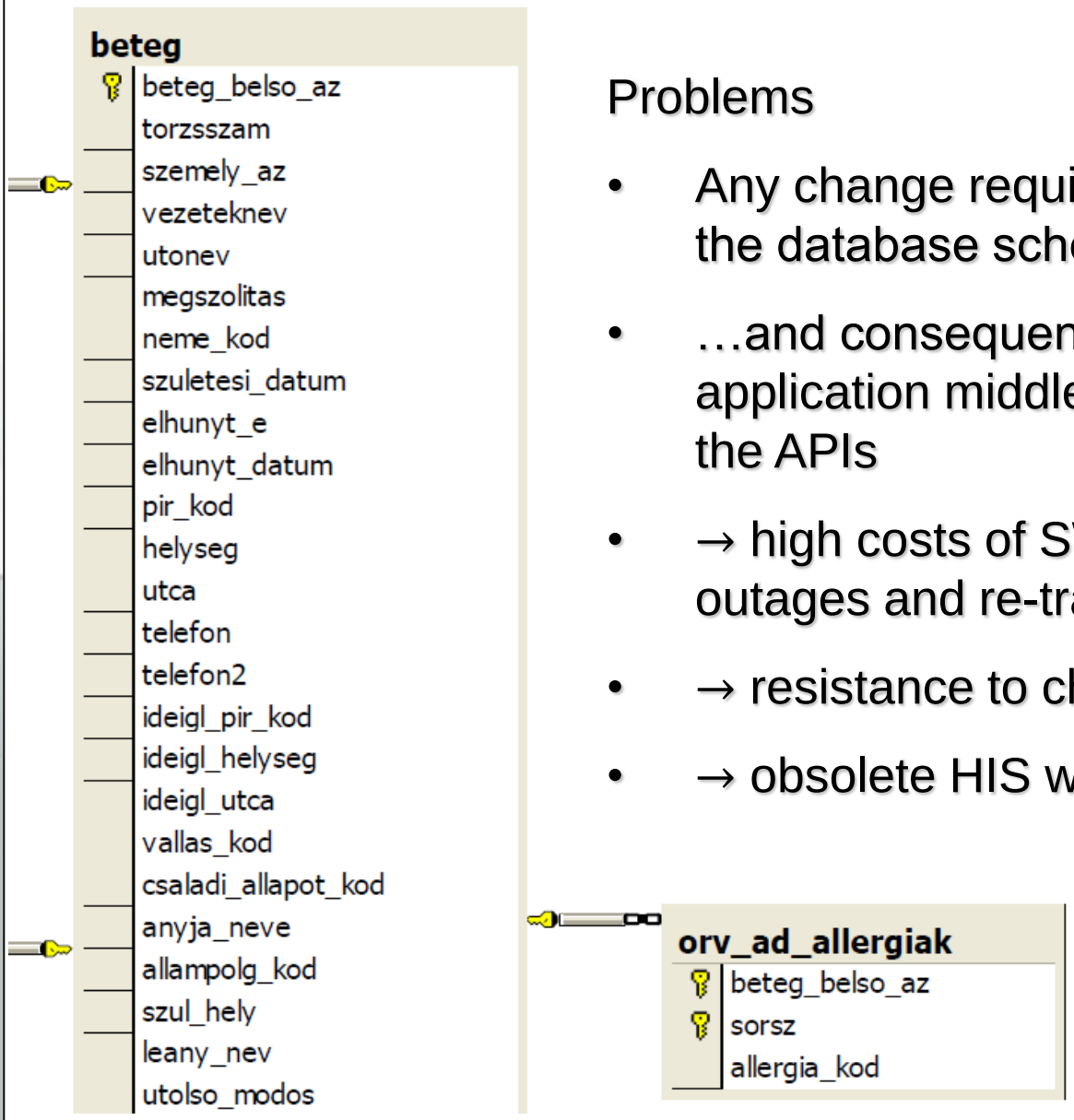
26. DIAGNÓZISOK

Megnevezés	T	Kód	D
-1	1		

27. BEAVATKOZÁSOK

Megnevezés	Intézet/osztály azonosítója	Szakmakód	Dátum	J	N	S	F	M	A	T	Kód	L	db
-1													

Problems with the “traditional” HIS data model



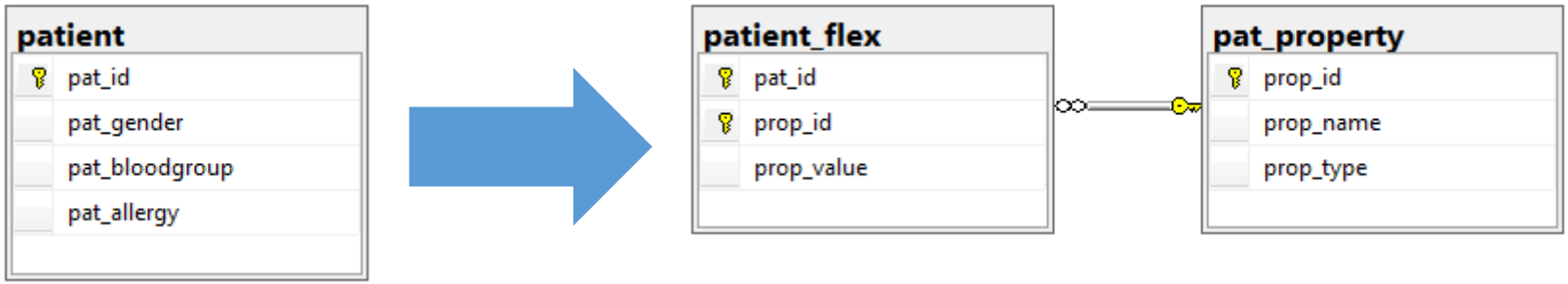
Problems

- Any change requires a change in the database schema…
- …and consequently in the application middleware or GUI, in the APIs
- → high costs of SW maintenance, outages and re-training
- → resistance to change
- → obsolete HIS with cryptic usage

A simple modeling technique

Let us make the data model more resistant to change:

- Use a **record-oriented** model instead of the **field-oriented** approach to resist changes in the *attribute set*



pat_id	pat_gender	pat_bloodgroup	pat_allergy
100	Female	*NULL*	Lidocain
101	Male	A+	*NULL*

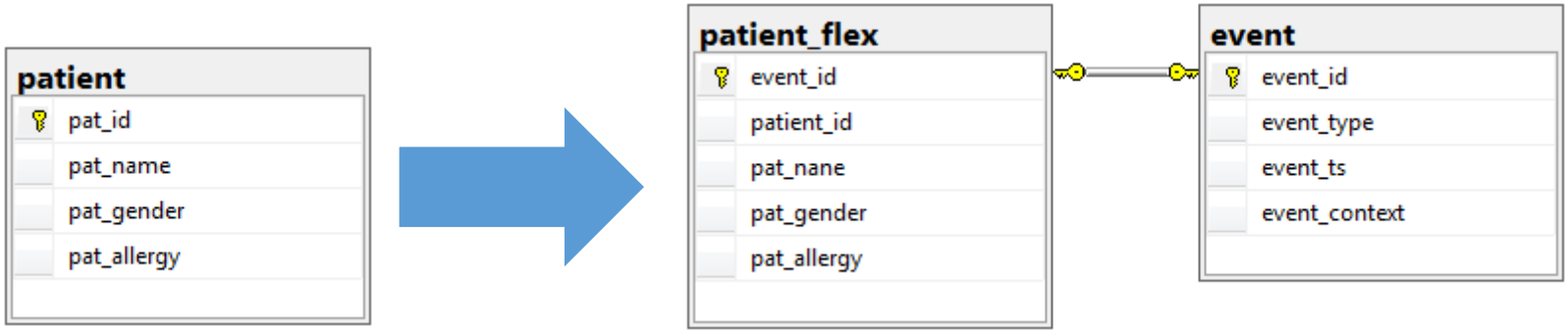
pat_id	prop_id	prop_value
100	1	Female
100	3	Lidocain
101	1	Male
101	2	A+

prop_id	prop_name	prop_type
1	gender	enum
2	blood group	enum
3	allergies	coded list

Another simple modeling technique

Let us make the data model more resistant to change:

- Use an **event-oriented** model instead of the **snapshot** approach to support changes *in time*



pat_id	pat_name	pat_gender	pat_allergy
100	Klein-Kiss Éva	F	Lidocain
101	Cser Péter	M	*NULL*

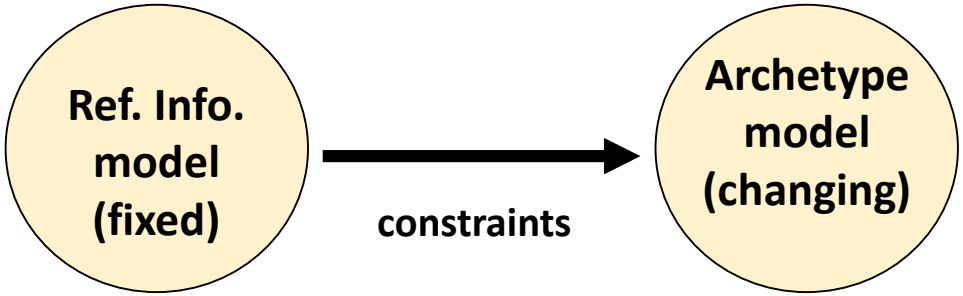
event_id	patient_id	pat_nane	pat_gender	pat_allergy
1	100	Kiss Éva	F	*NULL*
2	101	Cser Péter	M	*NULL*
3	100	Klein-Kiss Éva	F	Lidocain

event_id	event_type	event_ts	event_context
1	new pat.	2015-02-12	(...)
2	new pat.	2015-04-21	(...)
3	pat. update	2017-08-30	(...)

However, such models require more programming…

The two-level modeling concept

- The *structure* of data may also change. Simple techniques cannot support too complex data structures.
- The solution lies in between the medical and IT domains
- The idea is to *divide the domain model* into an IT model and a medical model
- Invented by the OpenEHR project, Thomas Beale, 1999



IT structures
(SW developer)
"grammar"

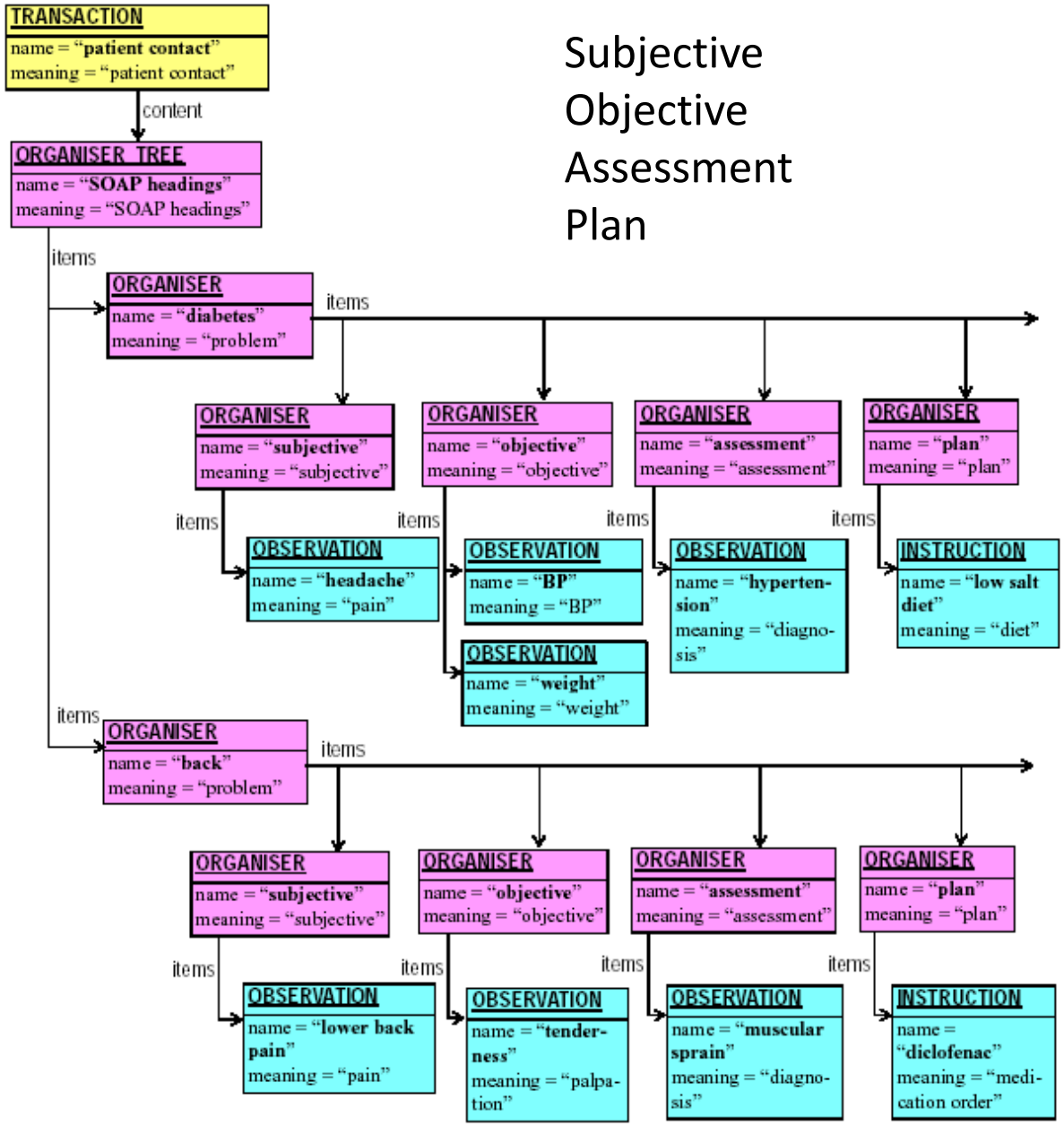
Medical structures
(Medical expert/authority)
"template"

The two-level concept simplified

The reference information model

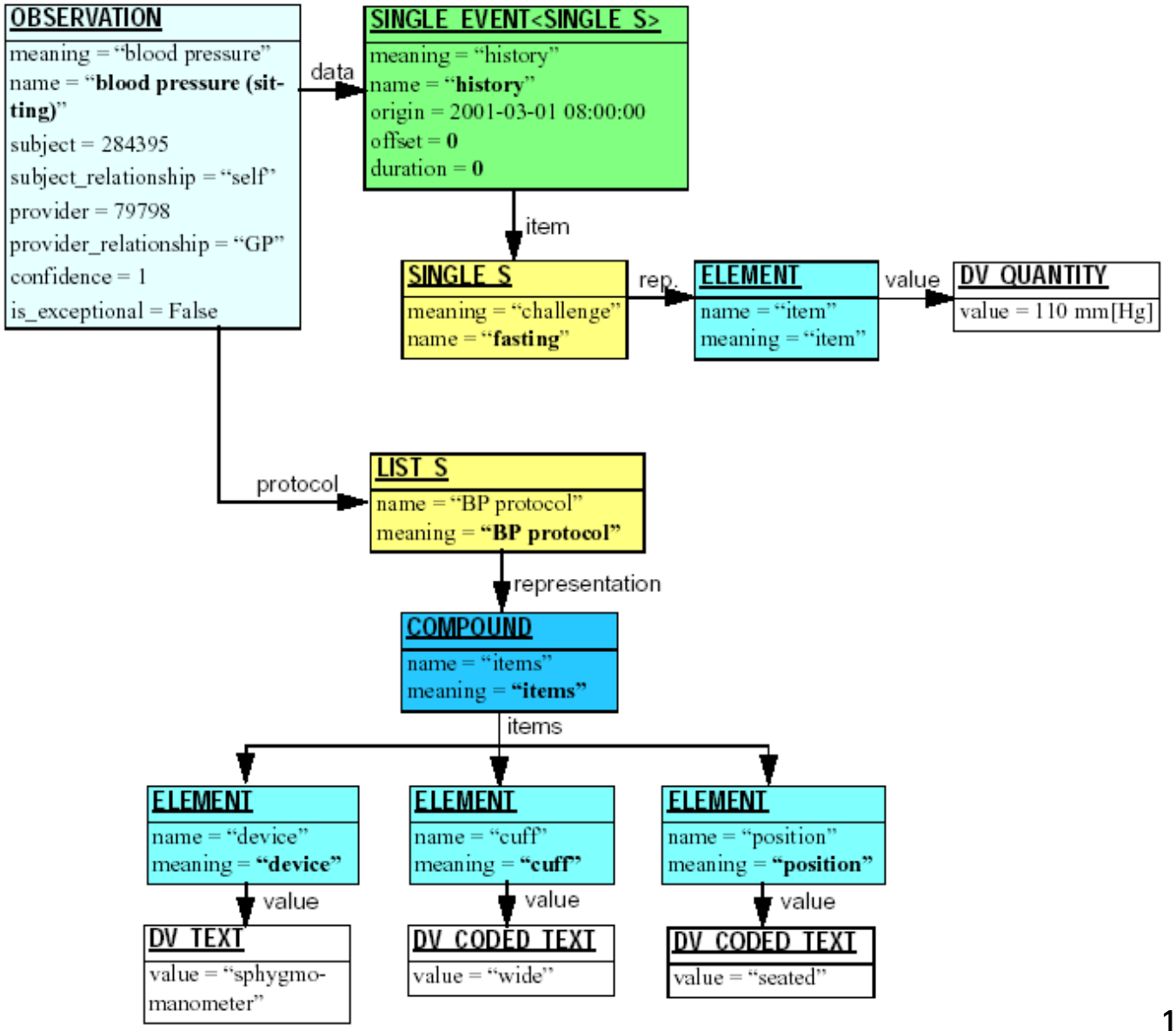
An archetype

SOAP: an openEHR archetype in use



Subjective
Objective
Assessment
Plan

The internals of an Observation in openEHR



An additional benefit: Interoperability

The two-level concept can be implemented as the native infrastructure of the HIS.

However, it also supports interoperability, with two levels of conformance:

- The two systems use only the same RIM.
  Contents can be parsed, checked, displayed, stored but cannot be semantically interpreted

  → *functional interoperability*

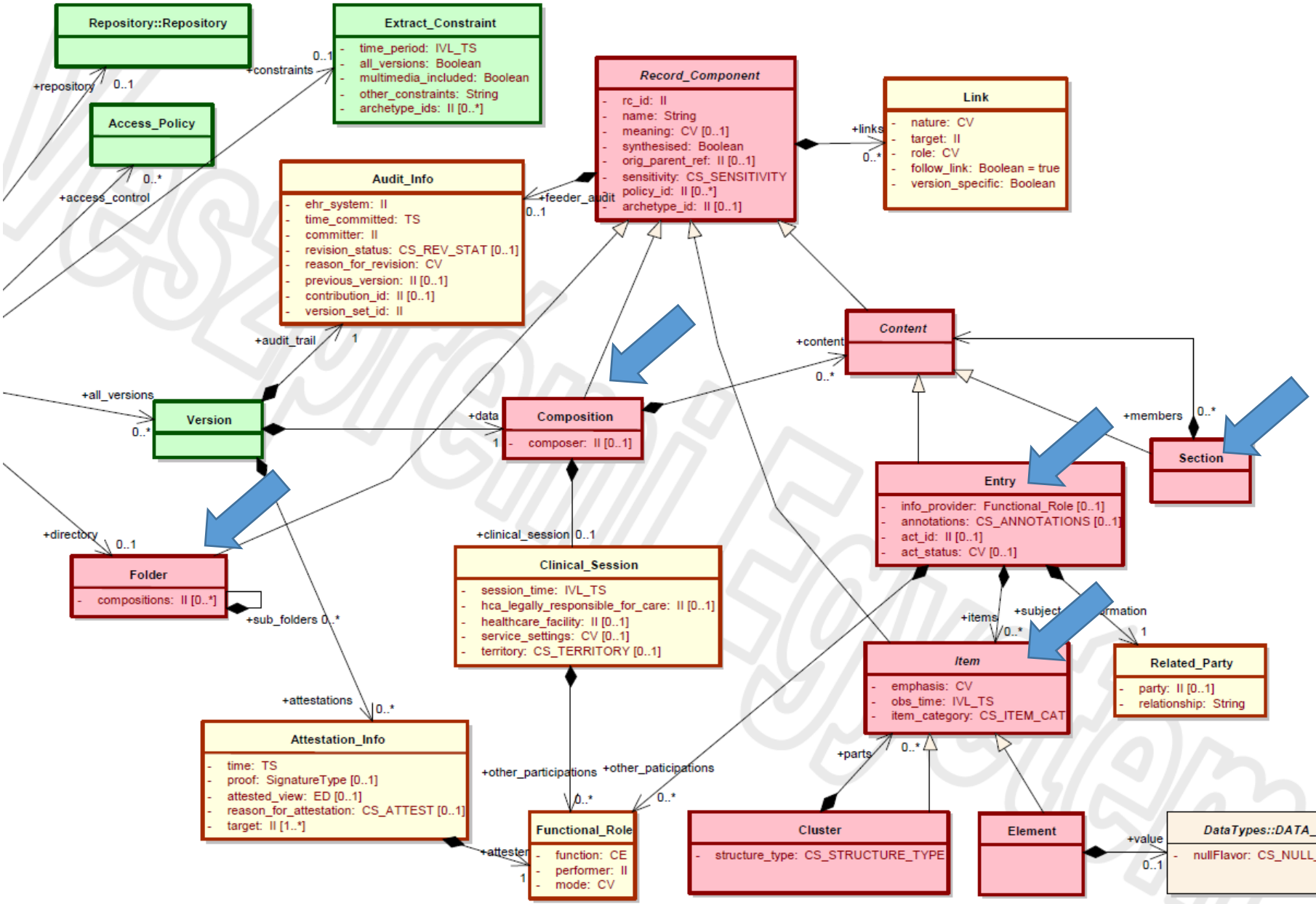
- The two systems also share the referenced archetype.
  Messages can be semantically processed and transformed into local data structures

  → *semantical interoperability:*
  communication can be fully automated without loss of information

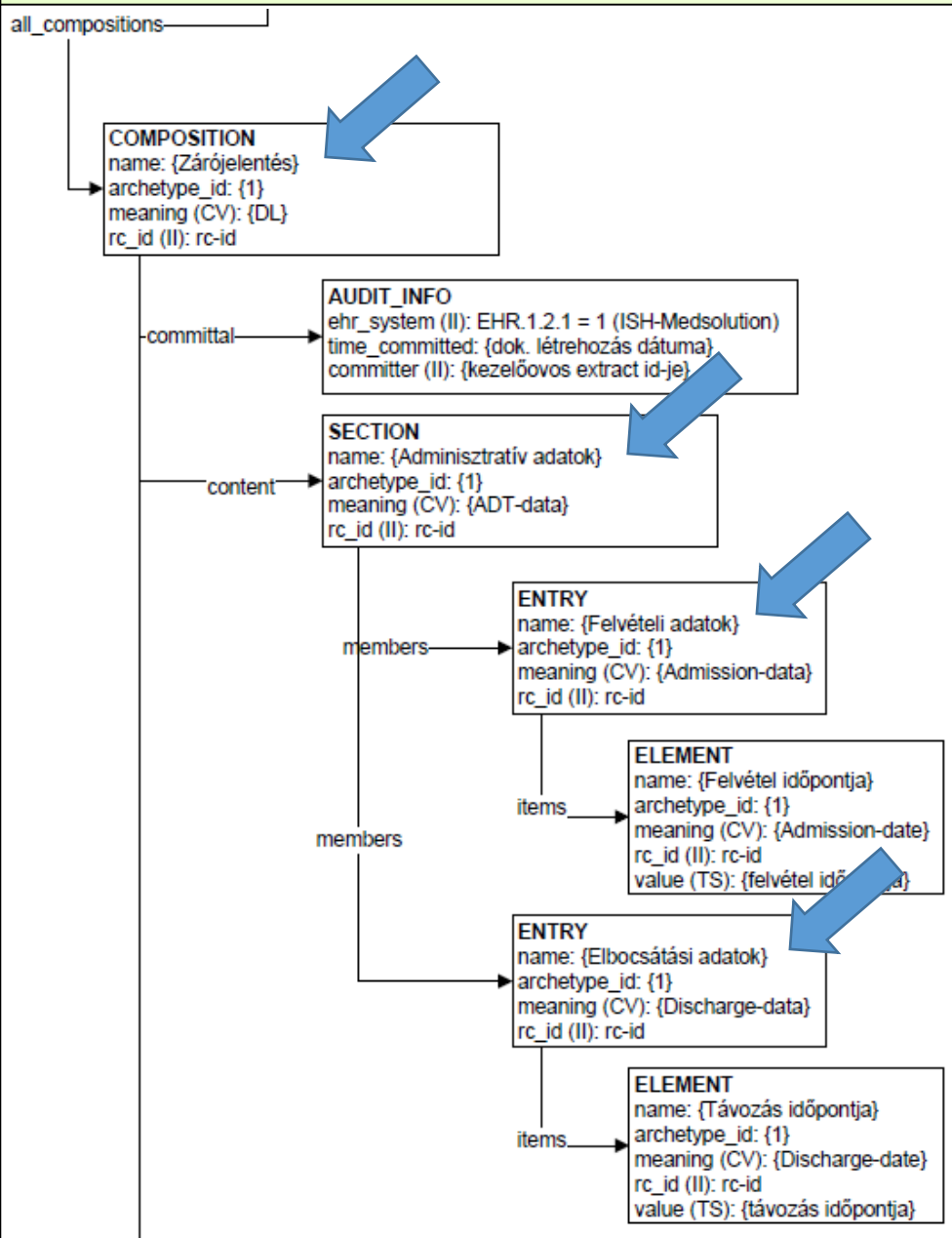
The two-level concept implemented in MSZ 22800

- A national implementation of the openEHR concept in 2004, adapted form CEN/ISO 13606
- Simplifications and extensions
- Main constituents of an EHR_EXTRACT RIM are
  - Compositions: events related to a single care event like an operation (versionable)
    - Sub-compositions and Sections for sub-topics
    - Legal and medical context
    - Entries in a tree structure, containing actual data in Elements
    - Elements can be clustered
  - Links connect any components e.g cause-effect
  - Folders: thematic collection of data e.g. ‘events related to the chronic diabetes of the patient’

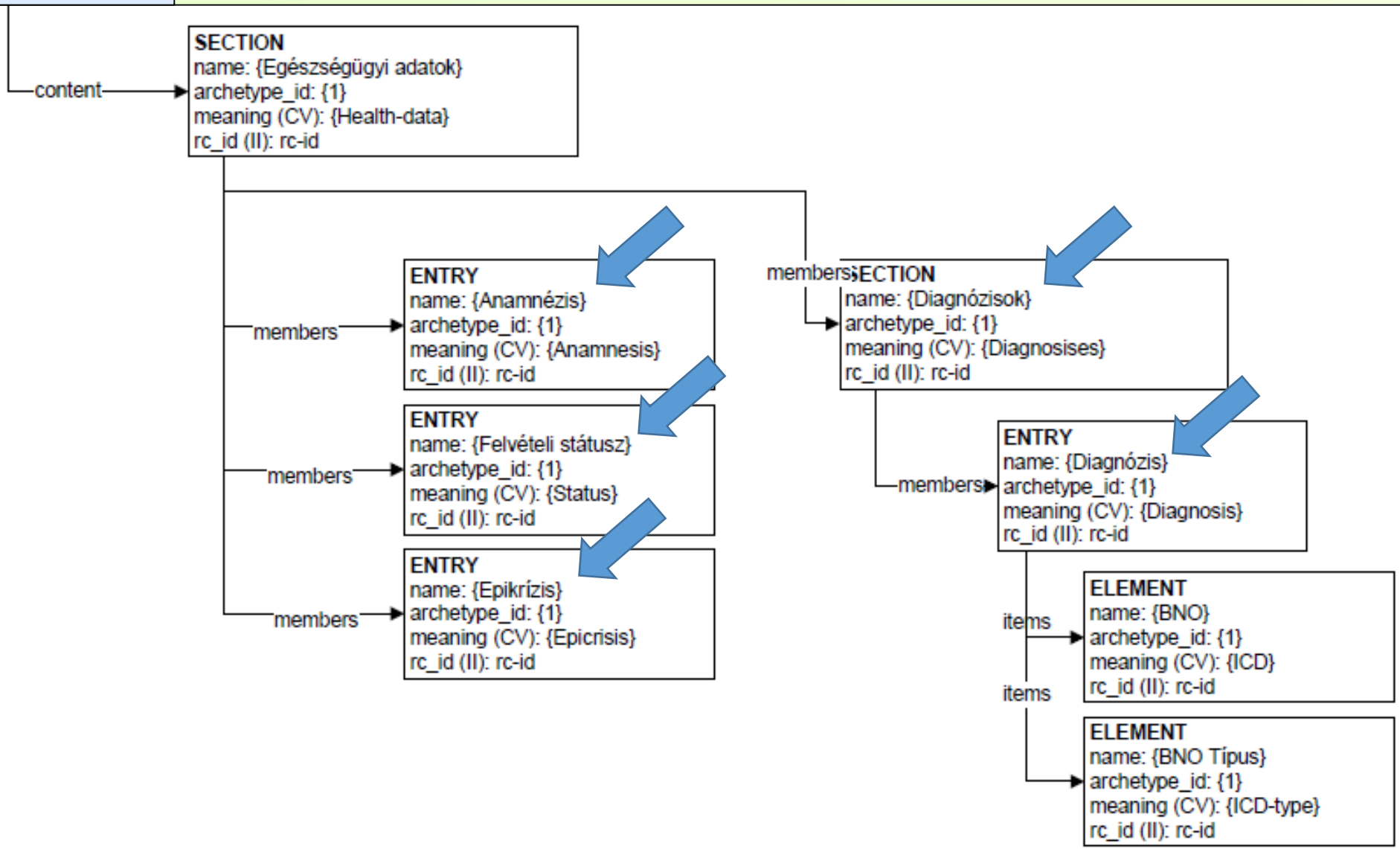
MSZ22800 Reference Information Model key entities



An MSZ22800 archetype for a discharge summary report



An MSZ22800 archetype for a discharge summary report



Implementing MSZ22800 structures in XML

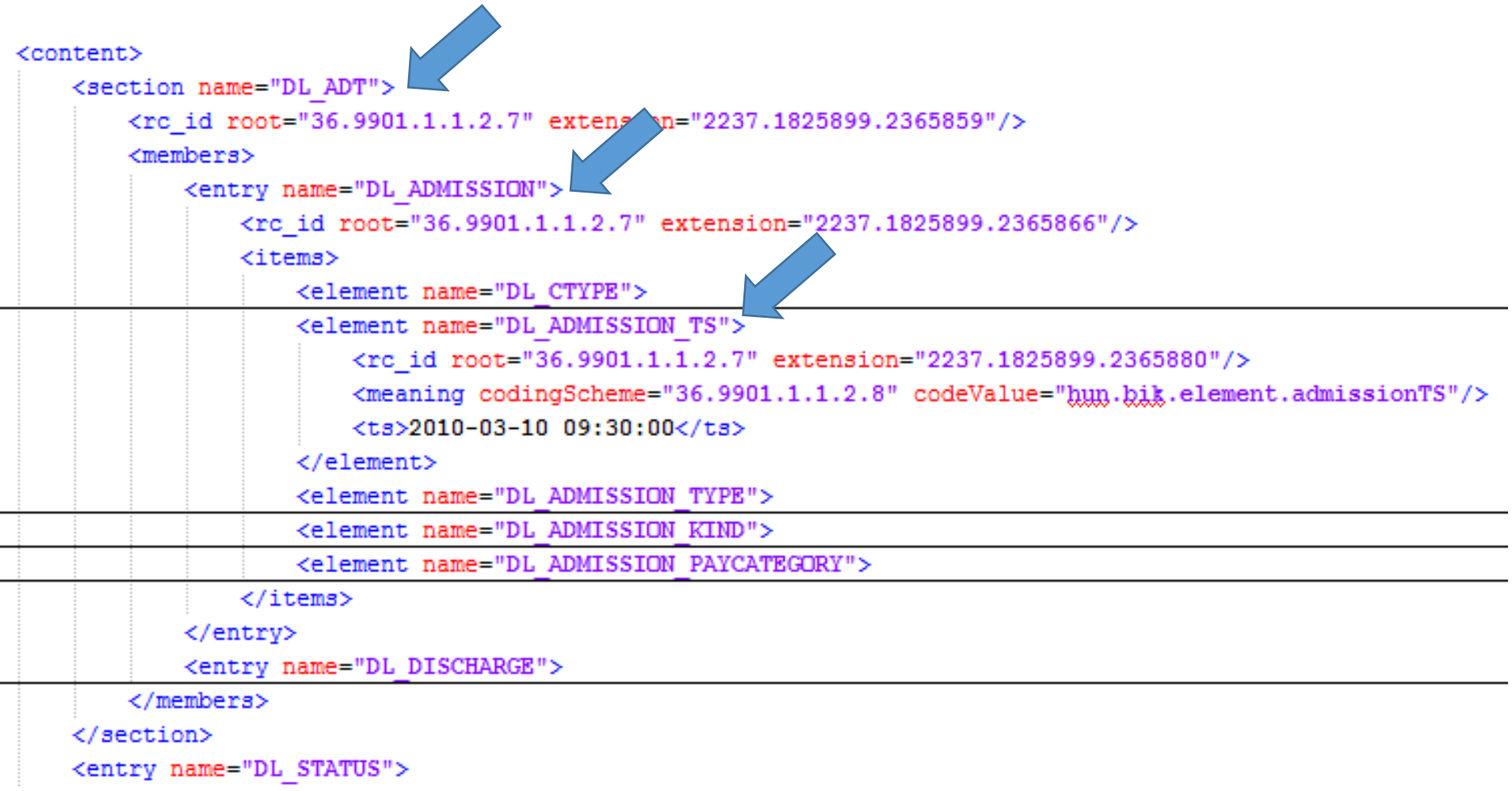
- XML is a proven, portable technology designed to support semantic interoperability
- W3C standard
- Built around the schema-validation concept (XML/XSD)
- XML is supported by all mainstream DB vendors as a searchable, updatable data type
- Conversion to/from relational structures and XSD management is a routine
- Also, native XML database engines exist

The validation of an EHR structure against an archetype can be formulated as the validation of a XML document against a schema

MSZ22800 document example in XML

```
<composition name="inpatientCase" archetype_id="hun-AT.bik.DL.v0">
    <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365852"/>
    <meaning codingScheme="36.9901.1.1.2.8" codeValue="hun.bik.composition.incase"/>
    <committal>
    <other_participations>
    <content>
        <section name="DL_DIAGNOSISES">
            <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365985"/>
            <members>
                <entry name="SYN_BNO">
                    <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365992"/>
                    <meaning codingScheme="36.9901.1.1.2.8" codeValue="hun.bik.content.diagnosis"/>
                    <items>
                        <cluster name="hun.bik.cluster.diagnosis" structure_type="ICD">
                            <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365999"/>
                            <parts>
                                <element name="SYN_ICDTYPE">
                                <element name="SYN_ICDCODE">
                                    <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2366013"/>
                                    <meaning codingScheme="36.9901.1.1.2.8" codeValue="hun.bik.element.icdCode"/>
                                    <string>M0590</string>
                                </element>
                            </parts>
                        </cluster>
                    </items>
```

MSZ22800 document example in XML

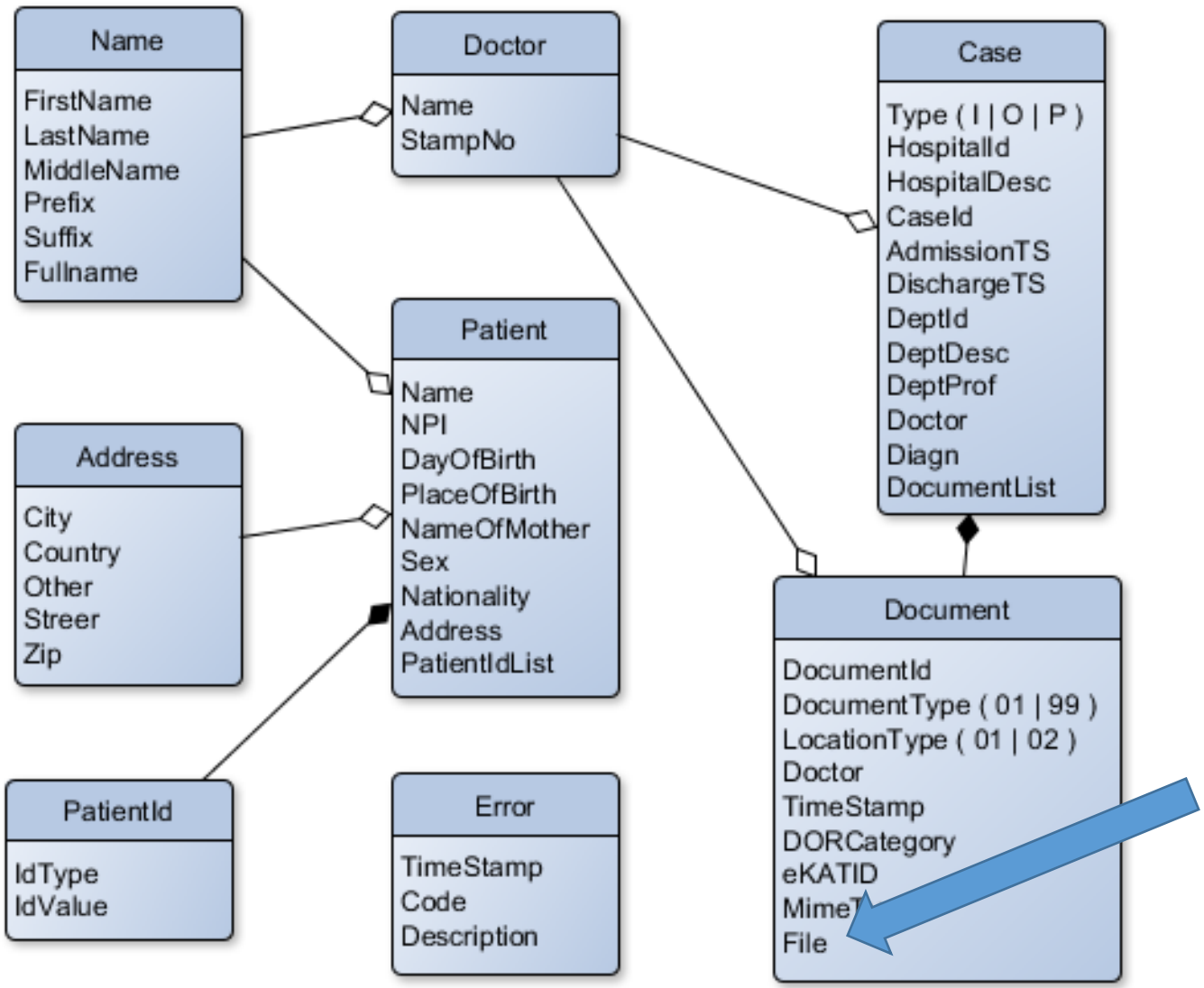


```
<content>
    <section name="DL_ADT">
        <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365859"/>
        <members>
            <entry name="DL_ADMISSION">
                <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365866"/>
                <items>
                    <element name="DL_CTYPE">
                    <element name="DL_ADMISSION_TS">
                        <rc_id root="36.9901.1.1.2.7" extension="2237.1825899.2365880"/>
                        <meaning codingScheme="36.9901.1.1.2.8" codeValue="hun.bik.element.admissionTS"/>
                        <ts>2010-03-10 09:30:00</ts>
                    </element>
                    <element name="DL_ADMISSION_TYPE">
                    <element name="DL_ADMISSION_KIND">
                    <element name="DL_ADMISSION_PAYCATEGORY">
                </items>
            </entry>
            <entry name="DL_DISCHARGE">
        </members>
    </section>
    <entry name="DL_STATUS">
```

The Electronic Health Cooperation Service Space (EESzT)

- A new national SOA information system for all actors of the health care domain
- Introduced gradually since 2013, in full power as of 1 November 2017
- All case reports to be uploaded in the central repository
- Uses the MSZ22800 as the domain model of the document in the report
- PDF is also allowed
- The first sub-domain to be implemented is e-Prescription, archetype ready…
- …but not yet applied

EESzT repository domain model



EESzT portal example

AEEK

ÁLLAMPOLGÁRI PORTÁL

FERENC FEKETE

EEGÉSZSÉGÜGY | TÖRZSEK ÉS NYILVÁNTARTÁSOK | JELENTÉSEK | ÜZENETEK | CSOPORTMUNKA | ÜZLETI NAPLÓ

TAJ szám 133066742 ÚJ KERESÉS VISSZA A FŐOLDALRA

Kezdő dátum 2015. 08. 31. Végdátum 2015. 09. 30.

SZŰRÉS

Műveletek MEGTEKINTÉS SZERKESZTÉS

1

Dokumentumok

○ Záró dokumentum (99), István Kiss (1234567), 2015. 09. 30.
Szent Margit Kórház (43), 2014. 06. 20. - 2014. 06. 20., Fül-, Orr-, Gégészet (111111111), Cholera vibrio (A0000)

Summary

- Modeling medical data is a complex task
- Continuous changes in attributes, values and structures are to be supported
- Simple techniques do not provide a robust solution
- We need the **two level modeling concept** based on a reference information model and specific archetypes
- Implemented in a national standard
- (will be) Applied in the new EESzT SOA

Thanks for your attention