



JOWO 2017

The Tyrolean Autumn - September 21 | 22 | 23

ODLS 2017 Ontologies & Data in Life Sciences

Basics of a drug ontology for annotations of clinical narratives

Zdenko Kasáč
Markus Kreuzthaler
Stefan Schulz

Institute for Medical Informatics,
Statistics and Documentation,
Medical University of Graz, Austria

Ontologies for Text Annotations

- Manually annotated texts: key resource for model-based information extraction
- Annotation scheme
 - clear-cut criteria
 - high reliability
 - no ambiguity
- Are ontologies good as annotation schemes
- Are their specific requirements / difficulties?

Use case: annotation of drug mentions

- Identification of drugs during a hospital stay crucial for quality assurance, detection of adverse events, retrospective studies, ...
- In many hospitals across Europe: No computerised drug prescription, only paper notes
- Mentions of drugs in progress notes, nursing documentation, discharge letters
- Difficulties
 - ambiguities, misspellings, constantly new drugs
 - mention of drugs in many different contexts

Example drug annotations in prescription section

Thrombo	Ass	100mg	0-1-0				
DrugName	DrugName	Strength	Regimen				
Sortis	80mg	0-0-1					
DrugName	Strength	Regimen					
Pantoloc	ret.	47,5mg	1	-	0	-	1
DrugName	DrugName	Strength	Regimen	Regimen	Regimen	Regimen	Regimen
Seloken	DA	2-0-0					
DrugName	DrugName	Regimen					
Oleovit	D3	1x	wöchentlich	(Do)			
DrugName	DrugName	Regimen	Regimen	Regimen			
Thyrex	10 mg	mg	1-0-0				
DrugName	Strength	Strength	Regimen				
Torasemid	1-0-1						
DrugSubstance	Regimen						
Antiflat	5mg	3	x	3	ml	bei	Blähungen
DrugName	Strength	Regimen	Regimen	Regimen	Regimen	Other	Other
Xatral	forte	0-0-0-1					
DrugName	DrugName	Regimen					
Dominal	25mg	1-0-0					
DrugName	Strength	Regimen					
Marcoumar	laut	Pass	bitte	um	Gerinnungskontrolle	beim	HA
DrugName	Other	Other	Other	Other	Other	Other	Other
Laevolac	3x2EL						
DrugName	Regimen						
Hypren	Kapseln	3x1					
DrugName	DoseForm	Regimen					
Bioflorin	500mg	p.	o.	1-0-0			
DrugName	Strength	Route	Route	Regimen			
Tavanic	40mg	s.c.	1x1	abends			
DrugName	Strength	Route	Regimen	Regimen			
Urosin	300	mg	dzt.	pausiert			
DrugName	Strength	Strength	Other	Other			

Typical Drug-related information in other document sections

- *Past history:*
 - "History of amphetamine abuse"
- *Allergies:*
 - "No known Penicillin allergy"
- *Evolution*
 - "after antibiotic treatment"
 - "Beta blockers suspended"
- *Lab findings*
 - "Carbamazepine serum level 5 $\mu\text{g/ml}$ "

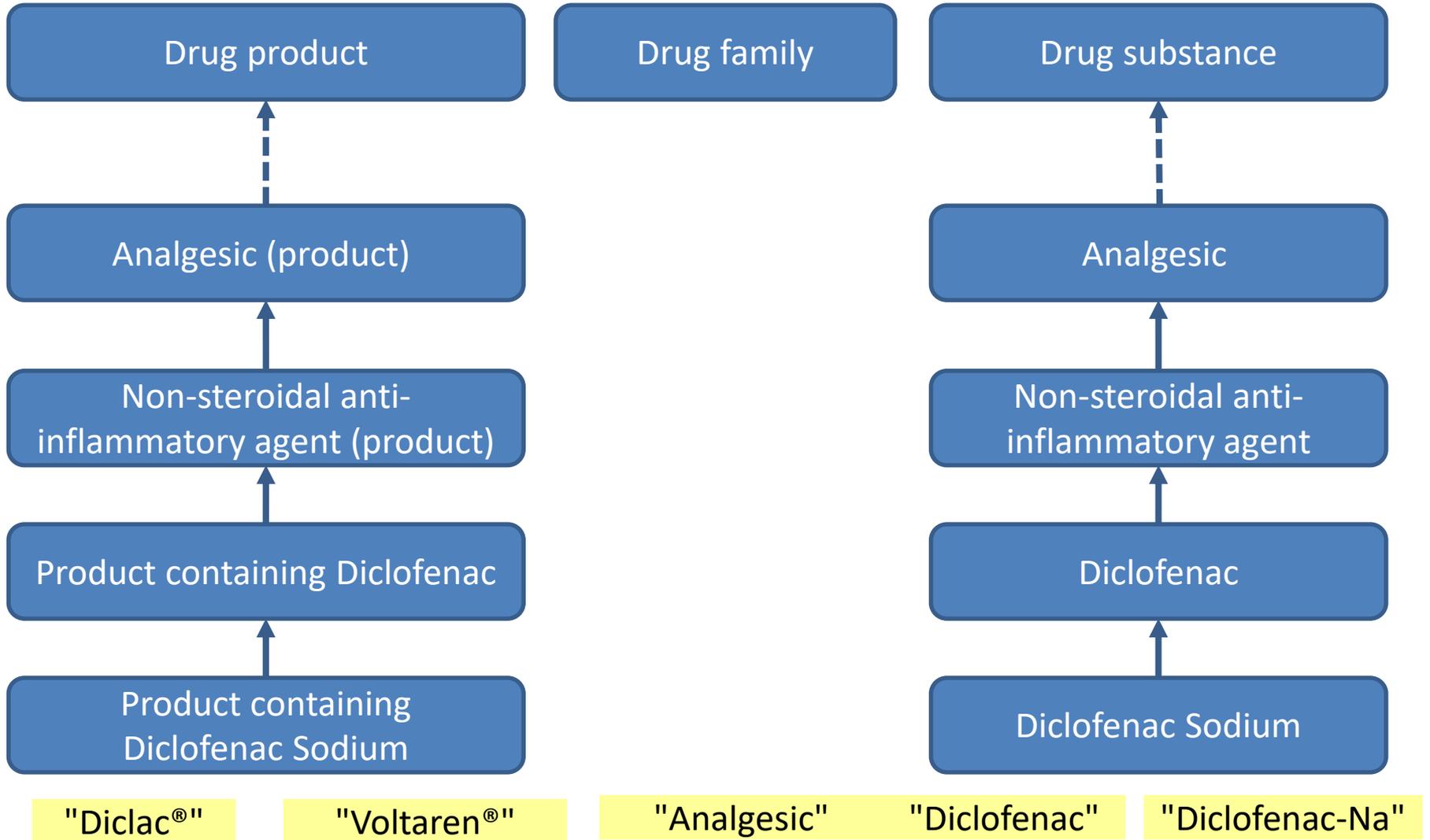
Dimensions of annotation

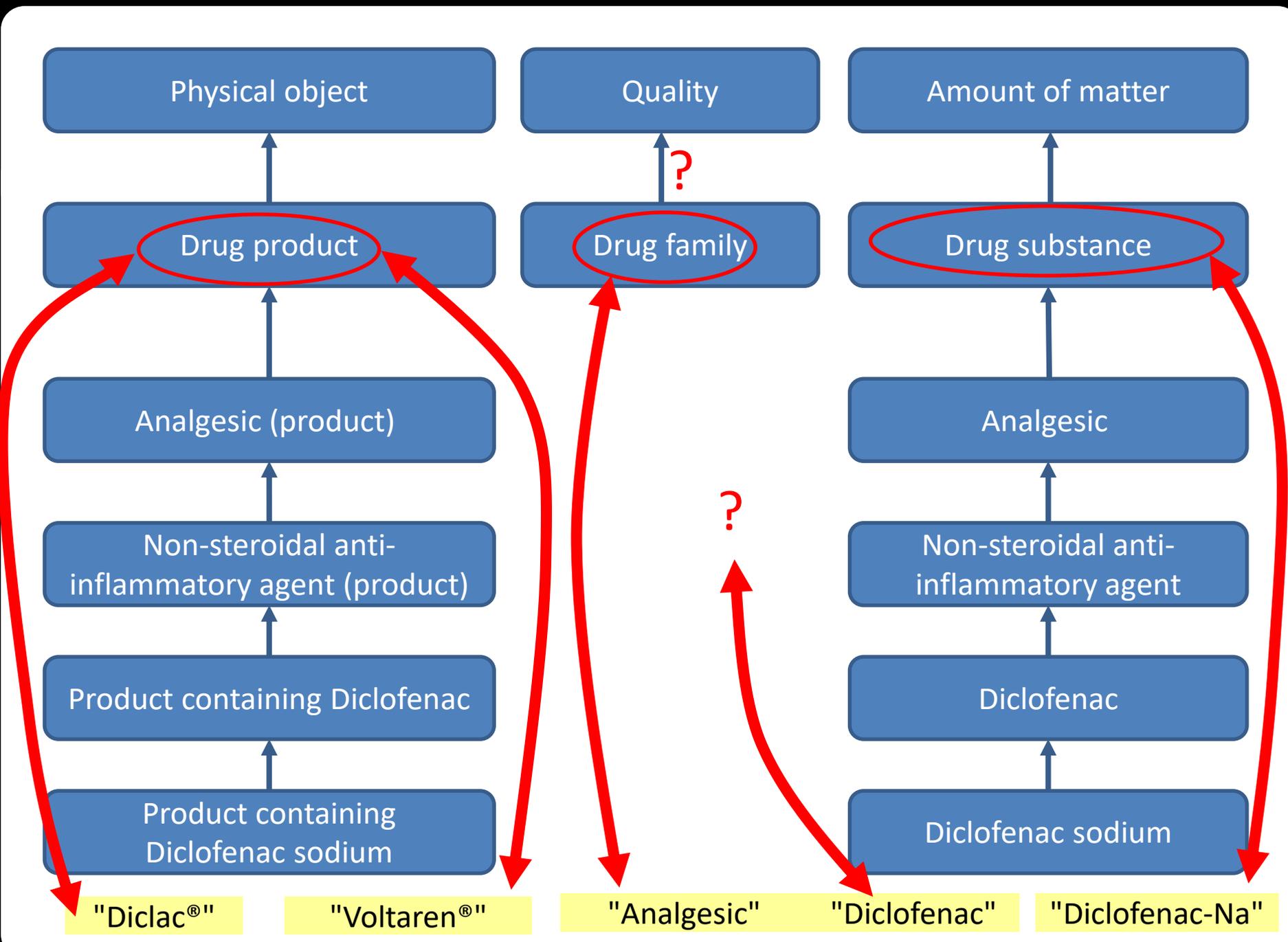
- Drug:
 - Drug substance, drug product, drug family, strength, dose form
- Administration:
 - Route ("oral"), regimen ("bd", "1-0-1"), other instructions ("before meal"), trigger conditions ("rectal temperature > 39°C)
- Context
 - current vs. past
 - prescribed, administered, suspended, changed
 - drug related risks, dispositions

Ontological issues: Drug

- Drug substance:
 - Amount of matter, mostly chemically defined, e.g. Diclofenac
- Drug product:
 - mostly industrially manufactured and registered, Brand name, e.g. Voltaren[®]
 - Drug products have drug substances as active ingredients
- Drug family
 - Non-terminal class in drug substance or drug product tree
 - Boundary issues

Example





Physical object

Quality

Amount of matter

Drug product

Drug family

Drug substance

Analgesic (product)

Analgesic

Non-steroidal anti-inflammatory agent (product)

Non-steroidal anti-inflammatory agent

Product containing Diclofenac

Diclofenac

Product containing Diclofenac sodium

Diclofenac sodium

"Diclac®"

"Voltaren®"

"Analgesic"

"Diclofenac"

"Diclofenac-Na"

Denotation

- Link between text sequence ("entity") and classes (concept) or logical expressions
 - Drugs, dose forms, regimen
- Context expressed by processual entity involved
 - administrating, prescribing etc.
- Epistemic status of denotation
 - e.g. if prescribed at discharge high likelihood that it has been given during hospital stay
 - drug allergy reported by patient vs.
 - drug allergy incident documented in EHR.

Denotation

- e Type (*DenotingEntity* and **bt12:represents** only
(*MedicationPrescription*
and **bt12:hasOutcome** some *DrugX*
and **bt12:hasOutcome** some *DoseFormX*
and **bt12:hasOutcome** some *RegimenX*
and ...))
- e Type (*DenotingEntity* and **bt12:hasPart** some *HighLikelyhood* and **bt12:represents** only
(*MedicationPrescription*
and **bt12:hasOutcome** some *DrugX* and ...))
- e Type (*DenotingEntity* and **bt12:represents** only
(*D* and **bt12:isIncludedIn** some *BloodSample*) and **bt12:isbearerof** some (*Concentration* and ...)

IDMP

Identification of Medicinal Products

IG 19844

Substances

ISO 11238

Data elements and structures for the unique identification and exchange of regulated information on substances

This norm distinguishes Substances (defined based on its main, general characteristics ; can have different roles e.g. active, adjuvant, basis of strength, excipient) and Specified Substances (More granular, specific description of a substance e.g. including manufacturing information, purity, grade ; allows for the specification of multiple substances "Intermediate Products" e.g. ASO3 - adjuvant composed of squalene (10.69 milligrams), DL- α -tocopherol (11.86 milligrams) and polysorbate 80 (4.86 milligrams))

IG 20440

Dose forms, etc.

ISO 11239

Data elements and structures for the unique identification and exchange of regulated information on pharmaceutical dose forms, units of presentation, routes of administration and packaging

Identifies for example injection solution, Injection suspension, Infusion solution (or a less granular regional term linked to these)

IG 20443

MPID

ISO 11615

Data elements and structures for the unique identification and exchange of regulated Medicinal Product information

Defines, characterizes and uniquely identifies regulated medicinal products for human use during their entire life cycle (Development, authorization, post-marketing and renewal or withdrawal from the market) ; Establishes definitions and concepts ; Describes data elements and their structural relationships required for the detailed description and unique identification of medicinal products

Units of measurement

ISO 11240

Data elements and structures for the unique identification and exchange of units of measurement

Specify rules for the usage of units of measurement for IDMP ; Define requirements for traceability to metrological standards ; Establish reference code system for units ; Provide structures and rules for mapping between different unit vocabularies and language translations, linking to existing systems, dictionaries and repositories

PHPID

ISO 11616

IG 20451

Data elements and structures for the unique identification and exchange of regulated Medicinal Product information

Pharmaceutical Product Identification (PHPID) based on the following subset of elements that describe the pharmaceutical product:

- Substance(s)/Specified Substance(s)
- Strength(s) - Strength units (units of measurement and/or unit of presentation)
- Reference Strengths
- Administrable Dose Form

