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
<table>
<thead>
<tr>
<th>DrugName</th>
<th>Strength</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
<th>Regimen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrombo</td>
<td>Ass</td>
<td>100mg</td>
<td>0-1-0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sortis</td>
<td></td>
<td>80mg</td>
<td>0-0-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pantoloc</td>
<td>ret.</td>
<td>47,5mg</td>
<td>1</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seloken</td>
<td>DA</td>
<td>2-0-0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oleovit</td>
<td>D3</td>
<td>1x</td>
<td>wöchentlich (Do)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyrex</td>
<td>10 mg</td>
<td>mg</td>
<td>1-0-0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torasemid</td>
<td>1-0-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antiflat</td>
<td>5mg</td>
<td>3</td>
<td>x</td>
<td>3</td>
<td>ml</td>
<td>bei</td>
<td>Blähungen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xatral</td>
<td></td>
<td>forte</td>
<td>0-0-0-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominal</td>
<td></td>
<td>25mg</td>
<td>1-0-0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marcoumar</td>
<td>laut</td>
<td>Pass</td>
<td>bitte</td>
<td>um</td>
<td>Gerinnungskontrolle</td>
<td>beim</td>
<td>HA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bioflorin</td>
<td></td>
<td>3x2EL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypren</td>
<td></td>
<td>Kapseln</td>
<td>3x1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bioflorin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tavamic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1x1</td>
<td>abends</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urosin</td>
<td>300</td>
<td>mg</td>
<td>dzt.</td>
<td>pausiert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DrugName</td>
<td>Strength</td>
<td>Route</td>
<td>Route</td>
<td>Regimen</td>
<td>Regimen</td>
<td>Regimen</td>
<td>Regimen</td>
<td>Regimen</td>
<td>Regimen</td>
</tr>
<tr>
<td>DrugName</td>
<td>Strength</td>
<td>Route</td>
<td>Regimen</td>
<td>Regimen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DrugName</td>
<td>Strength</td>
<td>Route</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DrugName</td>
<td>Strength</td>
<td>Route</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DrugName</td>
<td>Strength</td>
<td>Route</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Example drug annotations in prescription section*
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 μg/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
**Physical object**

- Drug product
  - Analgesic (product)
    - Non-steroidal anti-inflammatory agent (product)
      - Product containing Diclofenac
        - Product containing Diclofenac sodium
          - "Diclac®"
          - "Voltaren®"

**Quality**

- Drug family
  - Analgesic
    - Non-steroidal anti-inflammatory agent
      - Diclofenac
        - Diclofenac sodium
          - "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

\[ e \text{ Type } (\text{DenotingEntity} \text{ and } \\text{btl2:represents only (MedicationPrescription and } \text{btl2:hasOutcome some } DrugX \text{ and } \text{btl2:hasOutcome some } DoseFormX \text{ and } \text{btl2:hasOutcome some RegimenX and ... )})\]

\[ e \text{ Type } (\text{DenotingEntity} \text{ and } \text{btl2:hasPart some HighLikelyhood and } \text{btl2:represents only (MedicationPrescription and } \text{btl2:hasOutcome some } DrugX \text{ and ... )})\]

\[ e \text{ Type } (\text{DenotingEntity} \text{ and } \text{btl2:hasOutcome some } DrugX \text{ and ... } )\]

\[ e \text{ Type } (\text{btl2:represents only } (D \text{ and } \text{btl2:isIncludedIn some } BloodSample) \text{ and } \text{btl2:isbearerof some } (Concentration \text{ and ... }))\]
IDMP
Identification of Medicinal Products

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

- ISO 11238: Defines substances, their characteristics, and unique identifier
- ISO 11240: Specifies units of measurement
- ISO 11616: Defines pharmaceutical product information

Dose forms etc.

- ISO 11239: Describes dose forms and presentation

Source: ISO, 2012