

BACKGROUND KNOWLEDGE IN DATA MINING

Tomáš Klieger, Vojtěch Svátek, Milan Šimůnek, Daniel
Šťastný and Andrej Hazucha

BACKGROUND KNOWLEDGE

- Background Knowledge Exchange Format (BKEF) XML Schema
- Semantic abstraction over BKEF – the Background Knowledge Ontology (BKOn)

DESIGN OBJECTIVES

Disadvantages of semantic format as primary standard

- verbosity
- poor readability due to structural complexity
- the need for specialized, not widely available APIs

Proposed solution in form of XML Schema (BKEF)

DESIGN OBJECTIVES

Integration with Other Specifications

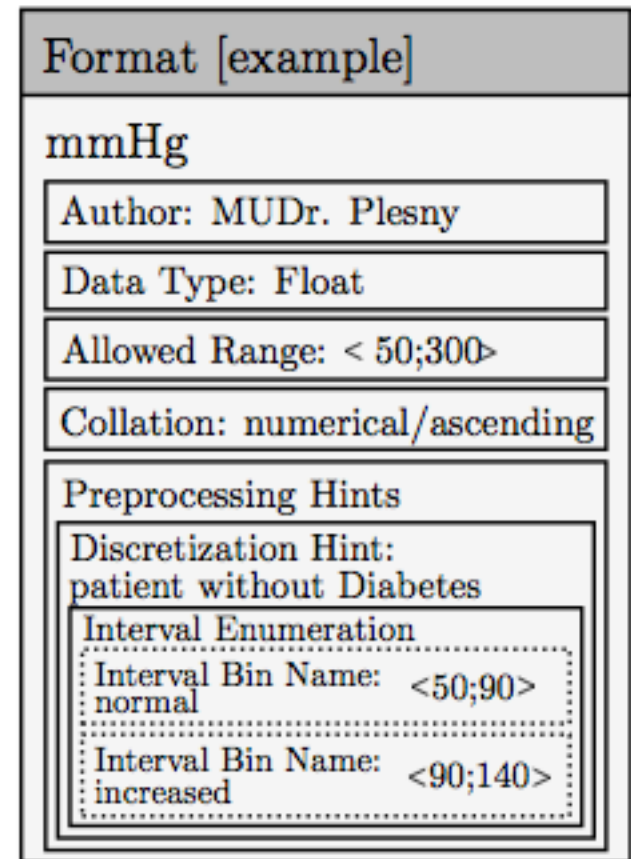
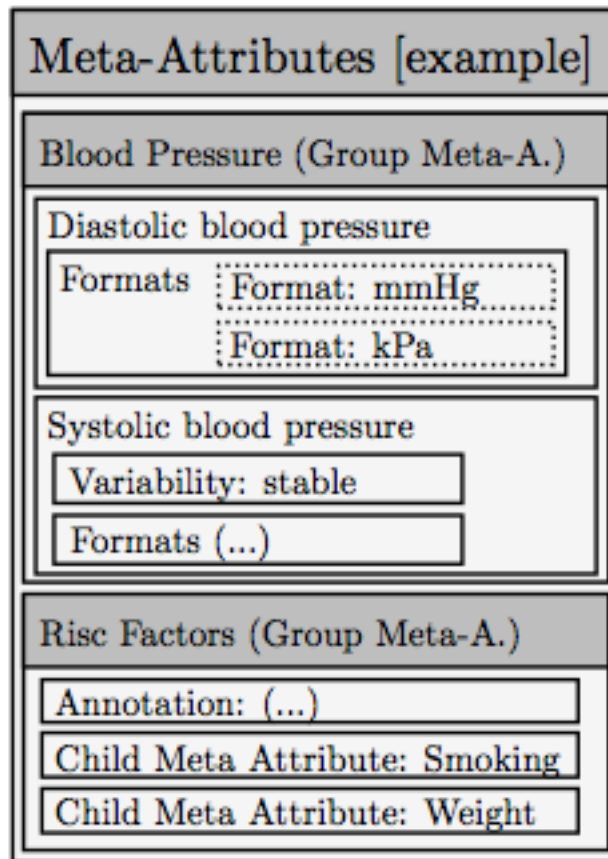
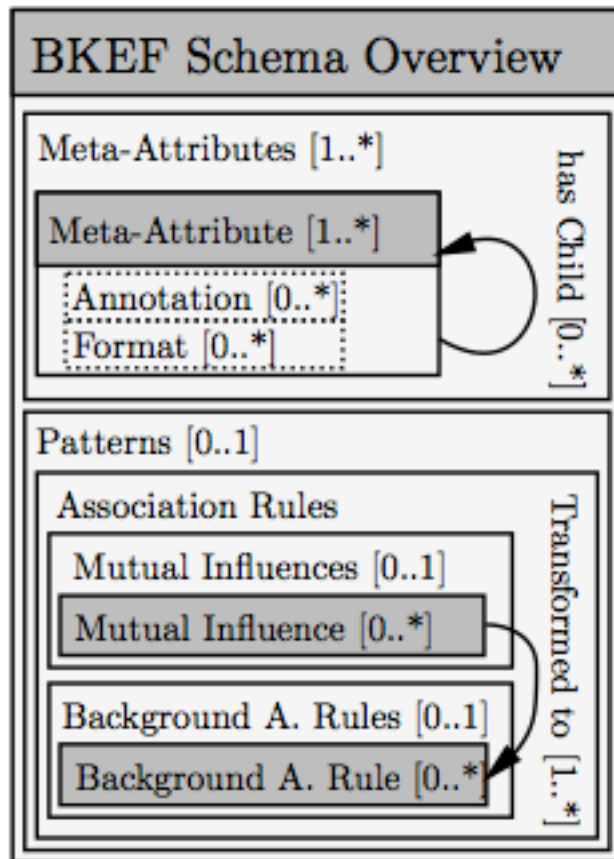
- PMML
- FML

BASIC CONCEPTS

- Hierarchically structured metaattributes
- Meta-field - combination of metaattribute and its format
- Meta-field Value
- Patterns - capturing relations between metaattributes

BKEF

stores mining models



BKOn

- semantic abstraction over BKEF
- adds typed associations between attributes
 - Semantic KB utilizes these relations for reasoning
- automatic transformation of BKEF XML into instances of the ontology concepts

BKEF, BKOn and Data Mining

- Data mining process initialization (Collecting knowledge from experts)
 - web based editor to build BKEF XML documents
- Linking Background Knowledge with Mined Data

BKEF, BKOn and Data Mining

- Background Knowledge for Localizing Search
- Background Knowledge for Result Pruning
- Background Knowledge for Postprocessing

CONCLUSIONS

- Standardized format for BK interchange in DM
- Ontology for BK modeling
- Future development: Ontology design patterns