

Harmonizing SNOMED CT with the BioTopLite Upper Level Ontology

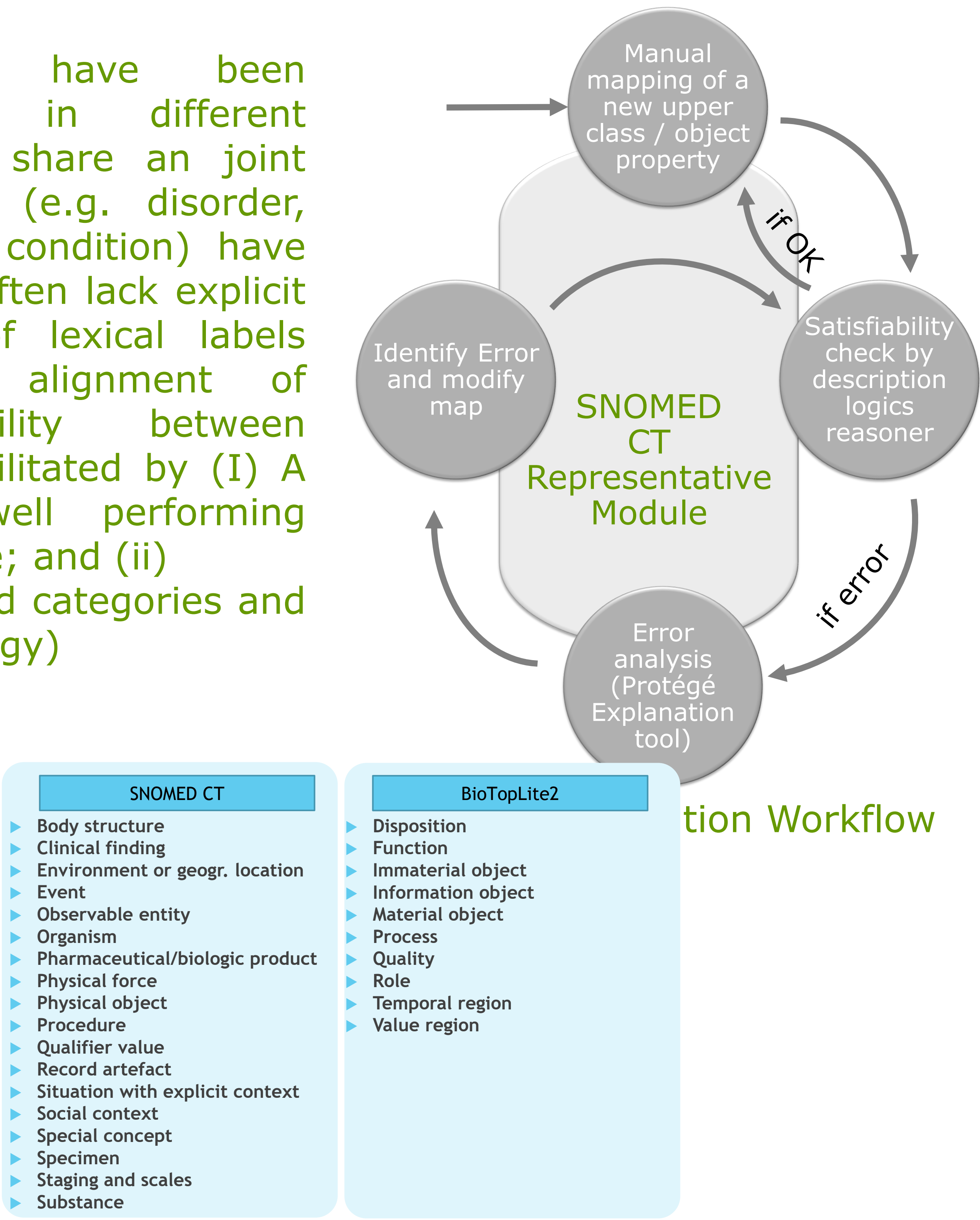
Stefan Schulz

Medical University of Graz (Austria)

Meaningful exchange of biomedical information requires support by ontologies and terminologies. SNOMED CT and others (e.g. OBO Foundry ontologies, WHO classifications) have recognized the need of precise descriptions of the entities denoted by terms and concepts, their ontological nature and the way they are related. They increasingly use a formal language, typically description logics (DL) axioms

Numerous ontologies have been developed bottom-up in different contexts. They do not share an joint upper-level. Key terms (e.g. disorder, animal, drug, situation, condition) have different meanings and often lack explicit definitions. Alignment of lexical labels does not guarantee alignment of meaning. Interoperability between semantic artefacts is facilitated by (I) A well-understood and well performing representational language; and (ii) a top-level layer of shared categories and relations (Top level ontology)

External sources cover content that is outside the scope of SNOMED CT but considered necessary for the common ontology, e.g. gene names or new content needed for ICD-11 but still in the submission process for SNOMED CT.



tion Workflow

The goal is to analyse the ontological structure of the OWL version of SNOMED CT in terms of

- Upper level concepts (classes)
- Relations
- Constraints

A preliminary alignment is done with the Upper-Level Ontology BioTopLite2 (BTL2). Consistency and performance are checked and the feasibility of moving to a richer language are assessed.

A collection of **Non-Description Logics Entities** constitutes the repository for all linearizations. They are linked via queries to the

concepts in the common ontology. These queries represent the numerous exclusion rules in linearizations and define non-ontological groupers (headings).

Linearizations are those releases of ICD-11 which address specific use cases like mortality, morbidity, primary care, reimbursement or classifications for medical specialties. They are familiar to the user, as they incorporate the classical classification principles (single hierarchy, non-overlapping classes, exhaustive partitions). They are expressed as queries on the common ontology, and incorporate additional knowledge from the ICD-11 content model. Residuals (NEC – not elsewhere classified, NOS – not otherwise specified) are automatically generated at all hierarchical levels. That linearizations are expressed by queries highlights their status as a special kind of terminological artefacts, which are not ontologies but whose content can be traced back to a principled ontology, viz. the **Foundation Ontology**. The hierarchical makeup may differ between linearizations, as they reflect pragmatic preferences in the arrangement of classification codes. Linearizations can also be nested.