

Using language technology for SNOMED CT localization



Stefan Schulz, Jose A Miñarro-Giménez

Medical University of Graz

Background

The absence of localized clinical terms mapped to SNOMED CT concepts constitutes a major barrier for SNOMED CT adoption. The translation of *Fully Specified Names* alone is debatable, due to the mismatch with the terms clinicians commonly use. The alternative is a bottom-up, incremental creation of interface terms connected to SNOMED CT IDs. Although this is not a translation in the first place it could also be used as raw input to speed up principled SNOMED CT translation efforts.

Examples for redundancy of sub-terms

(number of descriptions in which phrases occur) :

"Magnetic resonance imaging"	627
"off-road motor vehicle, except snow vehicle"	351
"second degree burn"	166

Examples for term similarities across languages (here English - German) :

"cryoglobulinemia"	"Kryoglobulinämie"
"cardioesophageal"	"kardioösophageal"
"chemothalamectomy"	"Chemothalamektomie"

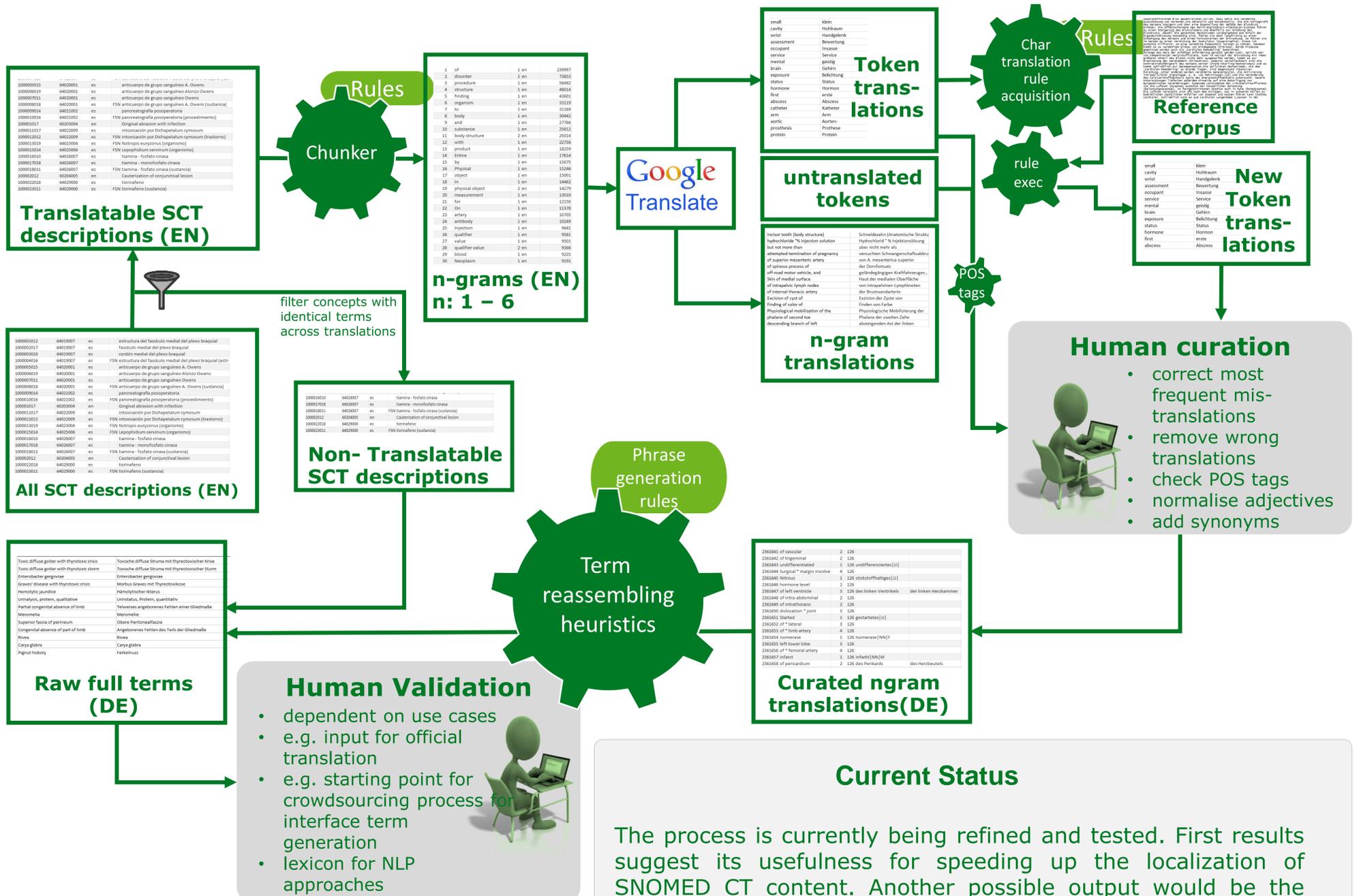
Web-based machine translation



Goal

A tentative workflow for semi-automated acquisition of interface terms linked to SNOMED CT is presented and described for the German language. It shows the potential of semi-automated term generation methods capitalizing on three factors:

- redundancy of SNOMED CT content
- word similarities
- Web-based machine translation.



Acknowledgements: European Commission SEMCARE – 7th FP grant 611388; ASSESS-CT - H2020-PHC-2014-15, grant 643818