Machine vs. Human Translation of SNOMED CT Terms

Stefan SCHULZ[a,b,], Johannes BERNHARDT-MELISCHNIG[a], Markus KREUZTHALER[a], Philipp DAUMKE[b] , Martin BOEKER[c]

[a]*Medical University of Graz, Austria*
[b]*University Medical Center Freiburg, Germany*
[c]*Averbis GmbH, Freiburg, Germany*

Background

- SNOMED CT: ontology-based, **international terminology** with over 300,000 concepts and over 700,000 English terms (Fully specified names (FSNs), preferred terms + synonyms)
- IHTSDO maintains English (US / UK) and (Latin American) Spanish version. English considered reference.
- Localised versions important for non-English speaking countries, creation and maintenance cost-intensive

Current translation projects

- Danish and Swedish versions completed (only FSNs)
- Ongoing translations: Canadian French
- Other European countries: translation of subsets
- Special situation for German speaking countries:
  - 2004 SNOMED CT completely translated by translation company, effort: 11.5 person years
  - Never released (copyright issues pending)
  - No IHTSDO member among the 7 countries in which German has the status of a primary or secondary official language

What should be translated?

Fully Specified Name (FSN)	Synonyms
Computerized axial tomography of brain (procedure)	Brain CT
Cerebrovascular accident (disorder)	CVA, Stroke
Sodium chloride solution (substance)	Saline, NaCl
Automobile, device (physical object)	Car

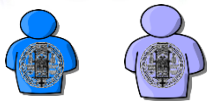
- Fully Specified Names (FSNs): standardized, self-explaining, lengthy
- Synonyms: represent clinical jargon, close-to-user, short, abbreviations, acronyms, ambiguous
- translations of **FSNs only** do not address important use cases (user friendly interfaces, natural language processing, layperson interfaces, …)

Alternative approaches

- Maintain English Fully Specified Names as ultimate reference for meaning (together with logical and (English) free text definitions
- Use low-cost translation methods for all terms (FSN, synonyms)
  - crowdsourcing targeting end users
  - non-expert translators
  - machine translation
- What about quality?

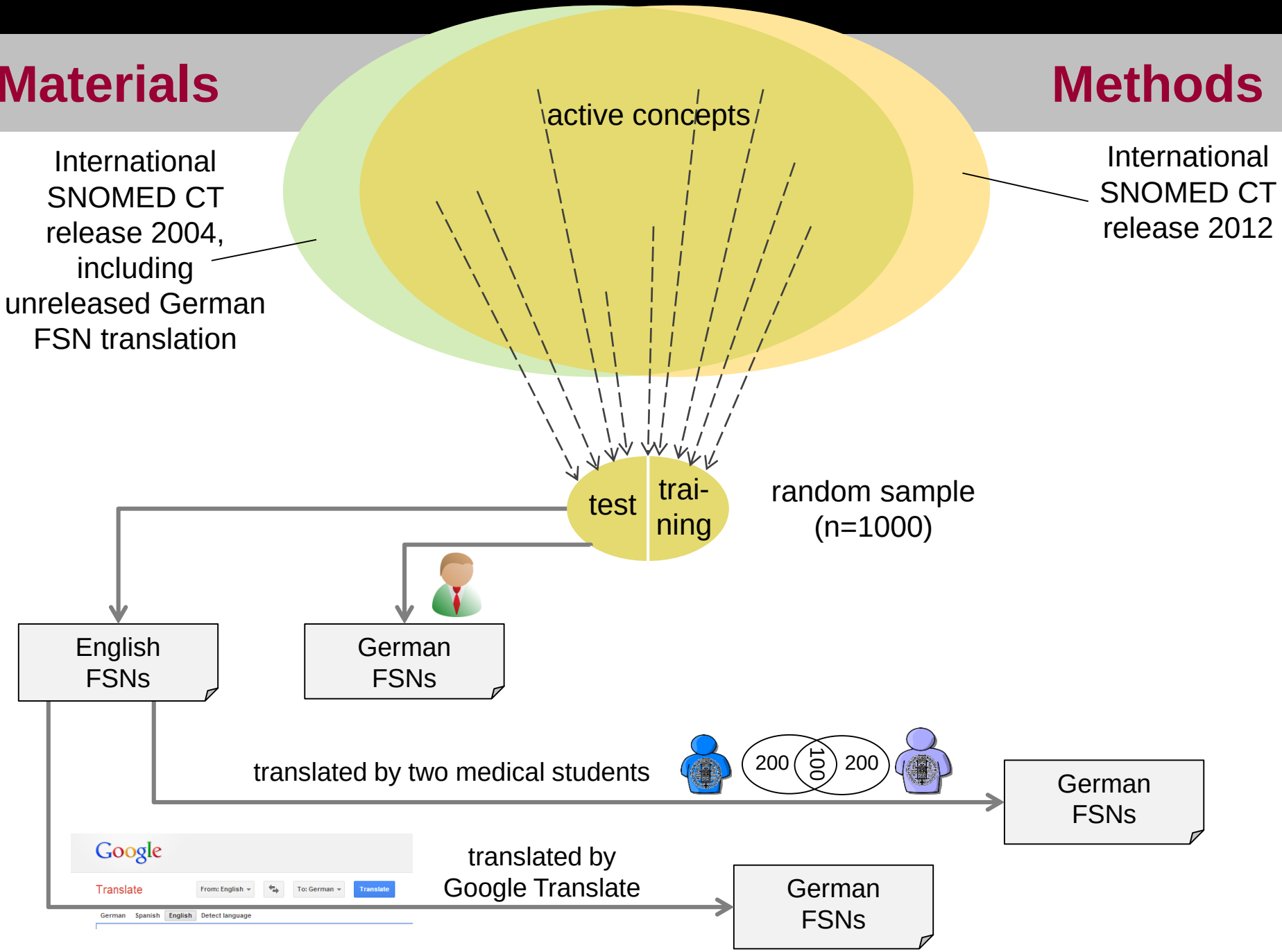
Study Objective

To compare three kinds of SNOMED CT translations from English to German

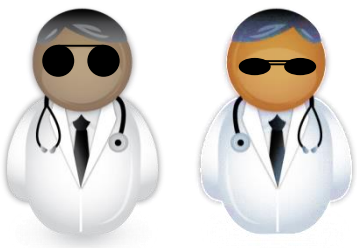
- Professional medical translators
- Free Web-based machine translation service Google Translate
- Medical students 

Materials

Methods



Scoring of the translations



- Blinded review by two domain experts

fully acceptable
marginally acceptable
unacceptable

- Semantic distance: modified Jaccard distance between sets of "semantic atoms" created by morphosemantic indexing *

fidelity of translation

linguistic correctness

Memantine hydrochloride 10mg tablet (product)							Memantinhydrochlorid, Tablette (10 mg)
Pseudoephedrine 30mg/5mL elixir (product)							Pseudoephedrin 30mg/5ml Elixier
Accidental netilmicin poisoning (disorder)							akzidentielle Netilmicinvergiftung
Eligibility for criminal injuries compensation (observable entity)		x		x			Berechtigung auf Kompensation für kriminelle Verletzungen
Moricizine hydrochloride (substance)							moricizine hydrochloride
Abdominal mass (finding)							Masse im Bauchraum
Entire retina of both eyes (body structure)							gesamte Retina beider Augen
Carboprost adverse reaction (disorder)							Carboprost-Nebenwirkung
Abnormal quantity (finding)	x			x			abnorme Quantität
Creation of external subdural drain (procedure)							Erstellung von externen subduralen Abfluss
Memantine hydrochloride 10mg tablet (product)							Memantinhydrochlorid 10 mg Tablette
Carcinoma in situ of jejunum (disorder)		x			x		Carcinoma in situ der Jejunum
Evoked potential, function (observable entity)		x			x		evozierten Potentials, Funktions
Injectable fibrinolysin (substance)							injizierbare Fibrinolysin
Hypopyon (disorder)	x			x			Hypopyon
Tongue symptoms (finding)	x			x			Zungensymptome
Myxedema neuropathy (disorder)	x			x			Myxödem-Neuropathie
Hydrocortisone valerate 0.2% ointment (product)							Hydrocortison Valerat 0,2% Salbe
Nurse: referred to (finding)		x		x			Krankenschwester: überwiesen
Collagen implant (substance)							Kollagen-Implantat
Fistulization of cisterna chyli (procedure)							Fistelung der Cisterna chyli
Gingival epithelial attachment (body structure)							gingival Epithelansatz
Massive silicotic fibrosis of lung (disorder)							Massive silikotische Lungenfibrose
Rev operation (procedure)							Rev-Operation
Ability to hit (observable entity)							Fähigkeit zu schlagen
Discharge by chemical pathologist (procedure)		x		x			Entlassung durch einen chemischen Pathologen
pM0 category (finding)							Stadium pM0
Malacia (morphologic abnormality)	x				x		malacia

* Daumke P, Schulz S, Müller ML, Dzeyk W, Prinzen L, Pacheco EJ. Subword-based semantic retrieval of clinical and bibliographic documents. Methods of Information in Medicine, 49:141–147, 2010.

Scoring Criteria

Es wird die sachliche und die sprachliche Korrektheit bewertet
Als externe Hilfsmittel sollen med. Wörterbücher, LEO und Wikipedia verwendet werden, ebenso wie (engl.) SNOMED -Browser

- Sachliche Korrektheit
  - Grün: Die Übersetzung gibt den Sachverhalt des Originals ohne Einschränkungen wieder, so dass sie zur klinischen Dateneingabe z.B. in Auswahllisten ohne Einschränkung verwendet werden können
  - Gelb: Die Übersetzung gibt den Sachverhalt des Originals mit Einschränkungen wieder. Für die Anwendung in der klinischen Dokumentation sollte die Übersetzung manuell überarbeitet werden
  - Rot: Die Übersetzung ist unbrauchbar.

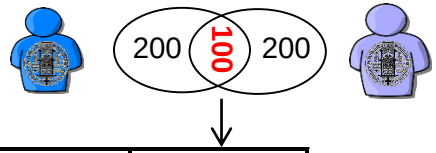
- Sprachliche Korrektheit: es wird rein der sprachliche Ausdruck unabhängig von der Übersetzung gewertet.
  - Grün: Die Übersetzung ist orthographisch und grammatisch einwandfrei, nach Vorgabe der von deutschen Medizinverlagen verwendeten Standards
  - Gelb: Die Übersetzung weist kleinere orthographische oder grammatische Mängel auf, die vor der Verwendung in klinischen Dokumenten korrigiert werden müssten
  - Rot: Die Übersetzung weist gravierende orthographische oder grammatische Mängel auf
  - Zu 2. Grün gehören kcz-Regel ("A. cerebralis", aber "Zerebralarterie"; "Ulcus ventriculi", aber "Magenulkus") Korrekte Verwendung von Bindestrichen, bzw. Zusammensetzungen (z.B. nicht "Antibiotika Therapie", sondern "Antibiotikatherapie" oder "Antibiotika-Therapie")korrekte Groß- und Kleinschreibung (am Termanfang optional)Die "Hierarchy Tags" (Klammerausdrücke) wurden bewusst nicht übersetzt"

Results: Translation

- Student translation performance (student translators)

90 sec / term → 6.3 person years for complete SNOMED CT

- Inter-translator agreement



Term concordance	Count
No synonymy	11
Close synonymy	20
Complete synonymy	25
Minor differences in spelling and punctuation	17
Verbatim agreement	27

Results: Quality of translation

- Inter-rater Reliability (Exact Fleiss' Kappa):

  Content fidelity: 0.24

  Linguistic correctness: 0.40

- Comparison of methods:

	t_1 Professional Translators	t_2 Google Translate	t_3 Medical Students
Linguistic Correctness	2.84 2.80 – 2.88	2.23 2.18 – 2.29	2.84 2.81 – 2.88
Content Fidelity	2.78 2.75 – 2.84	2.54 2.45 – 2.60	2.86 2.83 – 2.90
Semantic Proximity	0.45 0.43 – 0.48	0.52 0.49 – 0.55	0.49 0.45 – 0.51

3 fully acceptable
2 marginally acceptable
1 unacceptable

Summary of Outcome

- No difference between professional and untrained translators
- Automated term translation weaker especially regarding linguistic correctness (word endings, word order)
- In terms of term content fidelity automated term translation better than expected
- Inter-rater agreement low, particularly regarding content fidelity (despite preceding training phase)
- Semantic proximity lowest for professional translators (tendency towards more idiomatic translations?)

Limitations

- Small sample size, especially for stratifying the results along SNOMED CT semantic tags (disorders, procedures, substances, organisms etc.)
- Small number of raters does not represent the variety of medical professions
- Criteria for judging content correctness still too weak (despite commonly agreed rating guidelines prior to the experiment)

Final remarks

- Both lay translators and machine translations should be considered when translating SNOMED CT content
- Human review of machine translated content necessary
- According to expected level of consistency and quality (e.g. conformance with naming conventions), expert review also necessary for lay translations
- Interesting approach for harvesting synonyms or entry terms
- Results suggests feasibility for using a combined crowdsourcing / machine translation approach

Acknowledgements

- International Health Terminology Standards Development Organisation (IHTSDO) for the provision of the unreleased German SNOMED CT version

INTERNATIONAL HEALTH TERMINOLOGY STANDARDS DEVELOPMENT ORGANISATION