A Whole-Person Function Dictionary for the Mobility, Self-Care
and Domestic Life Domains: a Seedset Expansion Approach

Ayah Zirikly, Bart Desmet, Julia Porcino,
Jonathan Camacho Maldonado, Pei-Shu Ho, Rafael Jimenez Silva, Maryanne Sacco (*equal contribution)

Background
Whole-person functional limitations in the areas of mobility (MOB), self-care
and domestic life (SCDL) affect most individuals with disabilities.

Contributions
- First publicly available dictionaries of
  manually annotated terms related to
  MOB and SCDL, to help in extracting
  information relevant to disability.
  https://github.com/CC-RMD-EpiBio/terminologies
- Terminologies provide the
  associated ICF codes, and terms
  and their ICF codes have been
  verified by domain experts.

Expansion Methods
- Medical resources: using UMLS and
  ICF and SNOMED mappings
- Lexical resources: using VerbNet
  for terms in dictionary do
  clean and normalize term
  if first word of term is verb then
  get all synonyms of term
  add synonyms + term[1: len(term)] to
  dictionary where ICF code == term's
  ICF code

Results
Dataset
Patient is able to walk with the help of cane or another assistive device. He reports that he can take a bath and clean himself.

Seedset: annotated physical therapy notes from NIH Clinical center (450 MOB, 30 SCDL) and 250 consultative examination files for MOB and SCDL.

426 MOB and 298 SCDL terms

Expansion Methods
Manually annotated data
Seedset
Expansion methods
Medical resources
Domain-experts filtering
Lexical resources
Domain-experts filtering
Clustering-based
Final terminologies

Lexical resources: using VerbNet
expanded terms
in-domain terms: related to ICD and SNOMED
out-of-domain terms: wikipedia
nearest neighbor clustering in high dimensional space with cosine similarity

Challenges: multiword expressions and words that are not verbs.

Clustering and embeddings based

Motivation
- Lack of available dictionaries
- Share and develop common understanding and background about mobility action terms in the medical-related communities
- Improve many tasks’ performance, especially in the area of information extraction