

RP2 - Analysis of unstructured texts from publications

Research Question

- There is currently no German terminology on melanoma
- Is a partially automated creation of terminology for malignant melanoma possible?
- Which tools and resources can be used to create a German terminology?
- Which standardized terminologies/ontologies should be used to link the terms?
- How can the usefulness/correctness be evaluated?
- Are word embeddings and transformer architectures like BERT (Bidirectional Encoder Representations from Transformers) helpful in a multilingual context?
- Is the approach transferable to other diseases?
- Is the terminology useful at the point of care?

Solution approach

- Semi-automated creation of terminology with the support of machine learning methods (word embedding/transformer architectures)
 - Creation of German terminology with support of transformer architectures
 - Normalization and assignment to existing terminologies/ontologies (e.g. UMLS, RadLex)
 - Publication and maintenance by professional societies
 - Application example: Presentation of personalised summaries of inclusion/exclusion criteria for clinical trials
- ClinicalTrials.gov**
- Evaluation of usefulness by experts and in other dissertation projects of the GRK (e.g. FP1, FP4, FP10)

State of the art

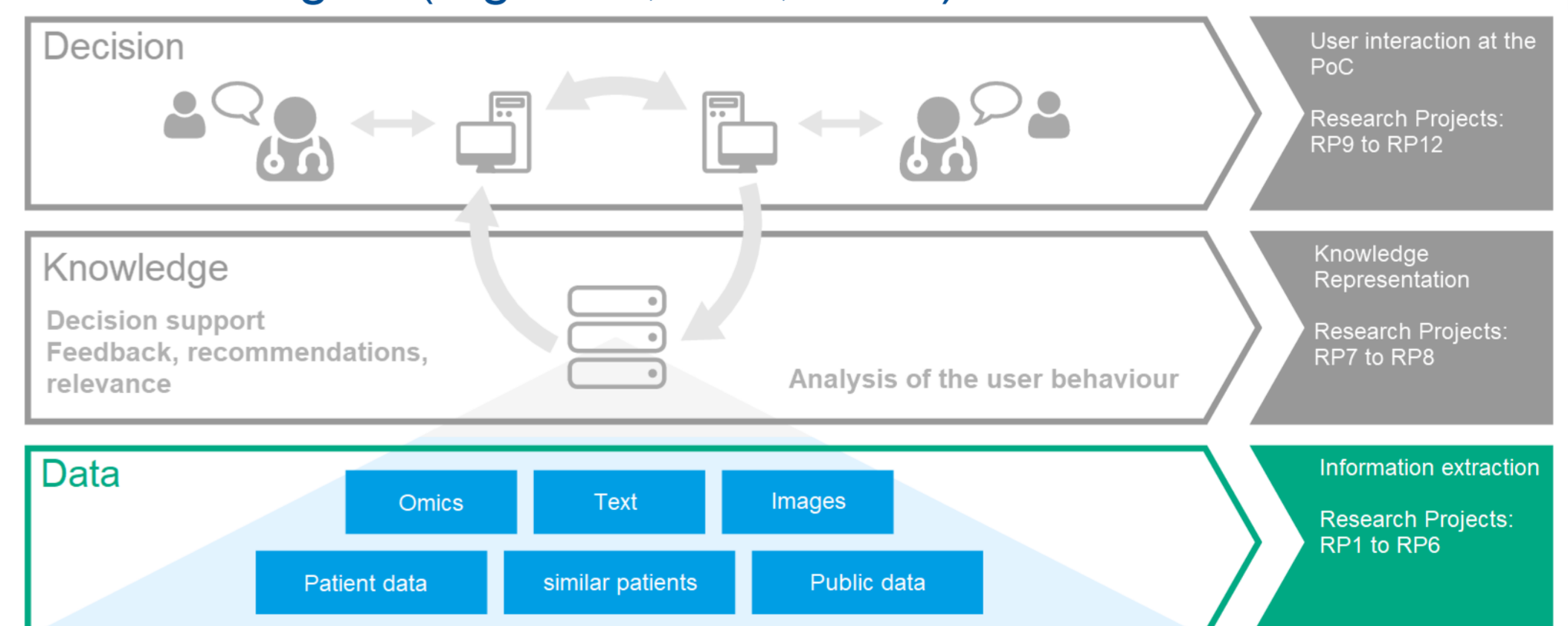
- Analysis of texts with English terminologies or ontologies, e.g. MaHCO [1], DermO [6] or ONTODerm [7]
- Terminologies not yet available in German, too unspecific and not comprehensive enough for applications
- Name recognition with cTakes or other systems and with machine learning possible [5].
- Partially automated creation for chemical name recognition and normalization possible [4].
- Name recognition for knowledge-based systems [8] and information discovery

Integration

- Support of terminology-marked text presentation

Primarius: **MM**, ED 03/2015, Lokalisation: **Rücken zentral**, TD 0,2mm, **keine Ulzeration**, **R0-Resektion**. Stadium bei ED: St. Ia pT1aNoM0 nach AJCC 2009
 2015: **R0 Resektion** extern, NE mit 1cm SA
 10/2017: Nachsorge über NFHA, dort **progrediente Schwellung der Achilla rechts**
 12/2017: MRT-Schädel: 4 mm RF rechts parietal subkortikal, V.a. **Metastase**, ED Stadium IV
 xx.12.2017 **Exzision** s.c. Ejlia OA proximal
 Ende 12/2017: PET-CT: Nachweis von **multiplen Metastasen: LK-Metastasen, muskuläre Metastasen, subkutane Metastasen** sowie **multifokale pulmonale Rundherde, Magenwand, DD Metastasen, RF Pankreas Kopf, RF Colon ascendens, mesenteriale, RF Tonsilla palatina rechts**
 xx.01.2018: Einleitung einer **Immuntherapie** mit **Nivolumab** (1mg/kg KG) und **Ipilimumab** (3mg/kg KG)
 xx.03.2018: 2. **Gabe Nivolumab** (1mg/kg KG) und **Ipilimumab** (3mg/kg KG)
 xx.03.2018 - xx.05.2018 **Transaminasenanstieg** Grad IV mit V.a. **Autoimmunhepatitis** Grad IV nach CTCAE
 xx.03.2018- 11.07.2018 **Pausierung Immuntherapie** mit **Nivolumab** (1mg/kg KG) und **Ipilimumab** (3mg/kg KG)
 xx.08.2018 Wiedereinleitung der **Immuntherapie** mit **Nivolumab** 480mg alle 4 Wochen bei steigendem S100 (0,22µg/l)
 xx.08.2018 - xx.10.2018 Z.n. 3 **Gaben Nivolumab - Monotherapie** 480mg
 10/2018 V.a. Immunermittele **Hypoxämie**
 xx.10.2018 Umstellung der **Systemtherapie** auf zielgerichtete mit **Dabrafenib** und **Trametinib** bei divergentem Ansprechen im letzten Stadium
 xx.11.2018 Einleitung **Dabrafenib** 150mg 1-0-1 und **Trametinib** 2mg 1-0-0
 Aktuell **Stable disease**

- Reference to explanations of terms/concepts from terminologies/ontologies
- Terminology/Ontology Server
- Support of other dissertation projects with the terminologies (e.g. FP1, FP4, FP10)



Prof. Dr. med.
Peter Horn

Universitätsklinikum Essen



Prof. Dr.-Ing.
Christoph M. Friedrich

Fachhochschule Dortmund

Literatur

1. DeLuca D.S., Beisswanger E., Wermter J., Horn P.A., Hahn U., Blasczyk R. (2009) MaHCO: an ontology of the major histocompatibility complex for immunoinformatic applications and text mining. *Bioinformatics*. 25(16):2064-70
2. Hahn U., Wermter J., Blasczyk R., Horn P.A. (2007): Text mining: powering the database revolution. *Nature*. 448(7150):130
3. Hahn U., Wermter J., DeLuca D.S., Bajwa A., Blasczyk R., Horn P.A. (2007) StemNet: An Evolving Service for Knowledge Networking in the Life Sciences; German e-Science Conference Proceedings, available online at <http://www.ges2007.de>
4. C. Kolarik, R. Klinger, C. M. Friedrich, M. Hofmann-Apitius und J. Fluck, „Chemical Names: Terminological Resources and Corpora Annotation“, Proceedings of the Workshop on Building and Evaluating resources for biomedical text mining (6th edition of the Language Resources and Evaluation Conference), Marrakech, Morocco, Seite 51-58, 2008.
5. R. Klinger, C. Kolarik, J. Fluck, M. Hofmann-Apitius und C. M. Friedrich, „Detection of IUPAC and IUPAC-like Chemical Names“, DOI:10.1093/bioinformatics/btn181, *Bioinformatics* 13(24), Seite i268-276, 2008
6. Fisher HM, Hoehndorf R, Bazelato BS, Dadras SS, King LE, Gkoutos GV, et al. DermO: an ontology for the description of dermatologic disease. *J Biomed Semantics*. 2016;7:38. doi:10.1186/s13326-016-0085-x.
7. Eapen B. ONTODerm - A domain ontology for dermatology. *Dermatology online journal*. 2008;14:16.
8. M. Hofmann-Apitius, J. Fluck, L. I. Furlong, O. Fornes, C. Kolarik, S. Hanser, M. Boeker, S. Schulz, F. Sanz, R. Klinger, H.-T. Mevissen, T. Gattermayer, B. Oliva und C. M. Friedrich, „Knowledge Environments Representing Molecular Entities for the Virtual Physiological Human“, DOI:10.1098/rsta.2008.0099, *Philosophical Transactions of the Royal Society A*, 366(1878), Seite 3091-3110, September, 2008.