

**CBmed GmbH**  
**Austrian COMET K1 Center for Biomarker Research**

Programm: COMET – Competence Centers for Excellent Technologies

Förderlinie: COMET-Zentren (K1) 3. Call, 2. Förderungsperiode

Projekttyp: 1.20 Digital Biomarkers for Precision Medicine / multi-firm



## KLINISCHE DOKUMENTENVERARBEITUNG FÜR DIE BIOMARKER-FORSCHUNG

### DBM4PM GEWINNT STANDARDISIERTE ONKOLOGISCHE PATIENTENPROFILE FÜR DIE ENTSCHEIDUNGSFINDUNG IN DER PERSONALISIERTEN MEDIZIN.

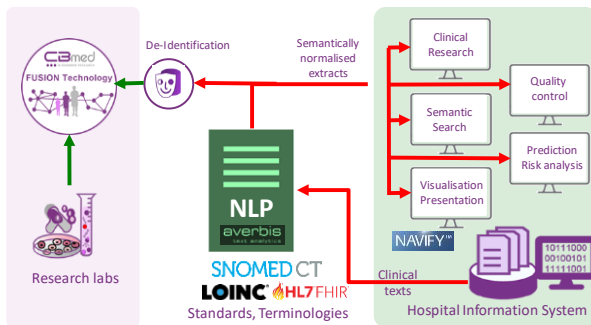
Das CBmed-Projekt DBM4PM (*Digital Biomarkers for Precision Medicine*) baut auf dem Projekt *Innovative Use of Information for Clinical Care and Biomarker Research* (IICCAB) aus der ersten CBmed-Förderperiode auf. Sein Ergebnis war eine klinische Dokumentenverarbeitungs-pipeline, die klinische Texte aus dem steirischen Krankenhausnetzwerk KAGes analysiert und aussagekräftige Informationen – wir nennen sie digitale Biomarker" (Erwähnung von Krankheiten, Phänotypen, Risikofaktoren, Scores, Medikamenten in der elektronischen Gesundheitsakte) – in standardisierter Form extrahiert. Die Verwendung internationaler Terminologiestandards wie SNOMED CT ist dabei ein wesentliches Alleinstellungsmerkmal.

Unsere Technologie – auch bekannt unter den Schlagworten NLP (*natural language processing*) und IE (*information extraction*) ist dadurch motiviert, dass ein großer Teil der klinischen Informationen nur als Freitexte verfügbar ist, während die Präzisionsmedizin immer mehr auf die

Wiederverwendung solch unstrukturierter klinischer Routedaten für Forschung und klinische Versorgung angewiesen ist.

Für die aktuelle Förderperiode hat sich Roche Diagnostics als Industriepartner dem Projektteam angeschlossen. Roche fungiert als Anbieter von NAVIFY®, einer Plattform für Tumorboards. In Tumorboards begutachtet ein Team aus Klinikern einen Tumorpatienten, um im Konsens die beste Behandlungsstrategie festzulegen. Der Erfolg von Tumorboards hängt entscheidend von einer benutzergerechten, strukturierten und selektierten Darstellung entscheidungsrelevanter klinischer Fakten ab.

## SUCCESS STORY



**Figure 1.** Klinische NLP (Natural Language Processing) als Service für Forschung und Entwicklung.

Im Gegensatz zu korrekturgelesenen medizinischer Literatur sind klinische Texte schwerer zu interpretieren, z.B. auf Grund von mehrdeutigen Akronymen und Abkürzungen, von Tippfehlern, Verneinungen oder nicht standardisierte numerischen Ausdrücken. Die klinische Sprache variiert darüber hinaus nach Institutionen, Berufsgruppen, klinischen Fachgebieten und geographischen Regionen. All diese Aspekte stellen daher besondere Herausforderungen für die Anpassung bestehender NLP-Lösungen oder die Implementierung spezialisierter IE-Module zur maschinellen Verarbeitung klinischer Textinhalte dar. Hier kommt unsere Technologie ins Spiel.

Im DBM4PM haben wir daher eine Kombination aus regelbasierten und maschinellen Lernmethoden für die ersten Schritte zu einem strukturierten

onkologischen Patientenprofil entwickelt. Entscheidend für beide Strategien ist der Zugang zu einer kritischen Masse an de-identifizierten und annotierten Trainingsdaten, um z.B. kontextualisierte Sprachmodelle für spezifische NLP und IE-Aufgaben zu erhalten. Durch die Nutzung der oben genannten Methoden in Kombination mit verschiedenen medizinischen Wissensquellen wie kontrollierten Vokabularien, Terminologien und Ontologien, insbesondere dem internationalen Terminologiestandard SNOMED CT, werden wiederverwendbare strukturierte Patientenprofilen angestrebt, die in Forschung und klinischer Versorgung vielfältig zu verwerten sind.

### Wirkungen und Effekte

Basierend auf der Arbeit der ersten Förderperiode wird nun eine auf deutschsprachige Kliniktexte optimiertes Inventar an IE-Komponenten weiterentwickelt und auf einen neuen Schwerpunkt ausgerichtet, nämlich die Aufbereitung klinischer Daten für Tumorboards. Diese Entwicklung erfolgt in enger Zusammenarbeit mit Roche DIS mit dem Ziel der Einbindung in NAVIFY®, um so Krebsbehandlung durch verbesserte klinische Entscheidungen zu optimieren. Die computergestützte Re-Strukturierung von klinischen Rohdaten ist ein wichtiger Schritt in Richtung der Anwendung der FAIR-Kriterien (Findable, Accessible, Interoperable and Reusable), auch über den Anwendungsfall Tumorboard hinaus

**CBmed GmbH**  
Stiftingtalstraße 5  
8010 Graz  
T +43 (0) 316 385 28801  
office@cbmed.at  
[www.cbmed.at](http://www.cbmed.at)

### Projektpartner

- SAP SE, Germany
- Roche Diagnostics GmbH, Austria

Diese Success Story wurde von der CBmed GmbH und den genannten Projektpartnern zur Veröffentlichung auf der FFG Website freigegeben. Das COMET-Zentrum CBmed GmbH, Austria COMET K1 Center for Biomarker Research / COMET-Projekt 1.20 wird im Rahmen von COMET – Competence Centers for Excellent Technologies durch BMK, BMDW, die Steirische Wirtschaftsförderung GmbH (SFG) und Wirtschaftsagentur Wien (WAW) gefördert. Das Programm COMET wird durch die FFG abgewickelt. Weitere Informationen zu COMET: [www.ffg.at/comet](http://www.ffg.at/comet)