

Recommender-System für Projektkollaborationen basierend auf wissenschaftlichen Publikationen und Patenten

PD Dr. Christoph Quix^{1,2}, Sandra Geisler², Rihan Hai²

¹ Fraunhofer-Institut für Angewandte Informationstechnik FIT
<http://fit.fraunhofer.de>

² RWTH Aachen University
<http://www.dbis.rwth-aachen.de>



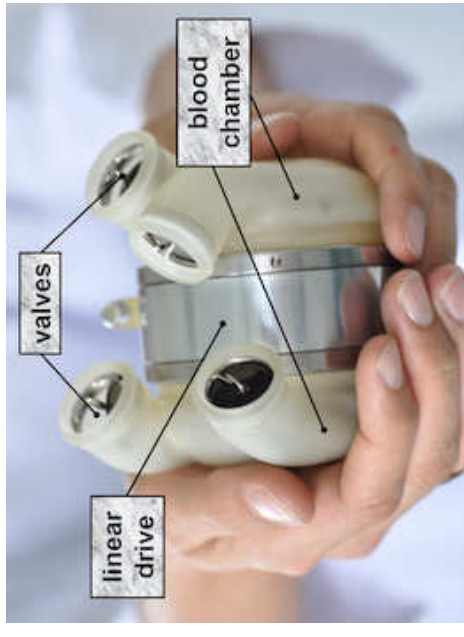
Prof. Dr. M. Jarke
Lehrstuhl Informatik 5
(Informationssysteme)
RWTH Aachen

Motivation

- ◆ Innovationen und Erkennen von neuen Trends sind entscheidend für Industrie & Forschung
- ◆ Medizintechnik
 - bedeutende, innovative Branche in Deutschland & weltweit
 - Interdisziplinär



Wikipedia



www.iem.rwth-aachen.de

Projektkollaboration

- ◆ Erfolg von innovativen Projekten häufig bestimmt durch Auswahl des Experten-Teams
- ◆ Partnersuche besonders in interdisziplinären Projekten schwierig, da „fremde“ Domänen mit einbezogen werden müssen

➔ Recommender-System zur Unterstützung bei der Suche nach Projektpartnern wünschenswert

Patente als Informationsquelle

- ◆ Erfinder sind meist produkt- und innovationsorientierte Experten
→ gut geeignet für Projekte mit schneller Markteinführung als Ziel
- ◆ Spezielle Sprache in den Patent erschwert Analyse

*The computer program is stored on a **computer-readable medium** comprising software code adapted to perform the steps of the **method 100** according to some embodiments when executed on a **data-processing apparatus**.*

mi-Mappa-Projekt

- ◆ Projektpartner für Medizintechnik-Projekte finden, basierend auf Patenten und andere produktbezogene Informationen
- ◆ Kooperation mit dem Institut für Medizintechnik der RWTH Aachen (AME)
- ◆ Gefördert durch
Laufzeit 2015-2017

Klaus Tschira Stiftung
gemeinnützige GmbH



<http://dbis.rwth-aachen.de/mi-Mappa>

Innovationsfelder

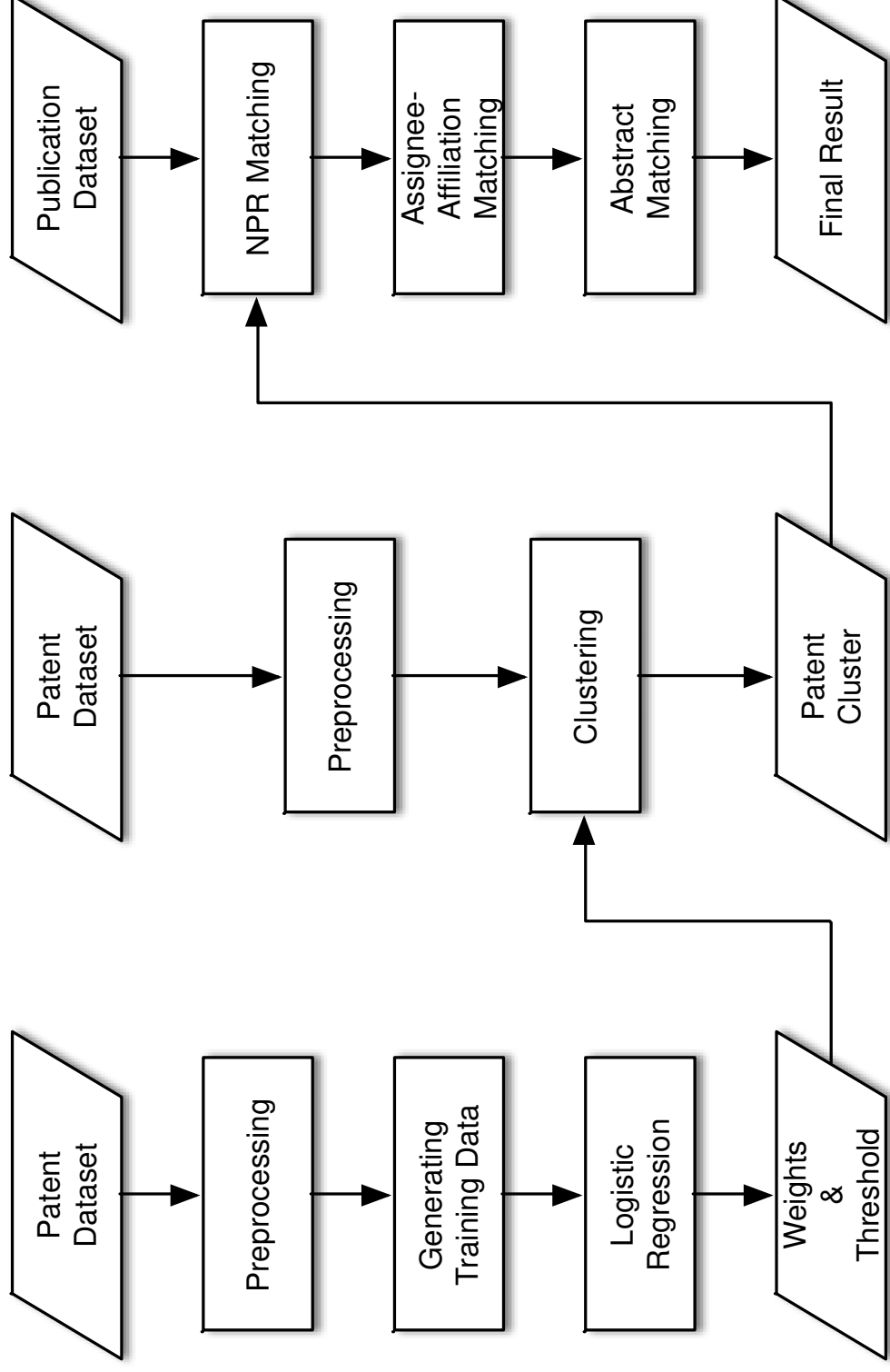
- ◆ Bereich mit signifikanten Innovationstätigkeiten
- ◆ Zukunftspotenzial
- ◆ Möglichst vollständige Wertschöpfungskette
- ◆ Definiert durch Expertengremium für die Medizintechnik [Schlötelburg et al., 2008]
- ◆ Beispiele:
 - Bildgebende Verfahren
 - Prothesen und Implantate

Lösungsansatz

- ◆ Experten auf Innovationsefelder abbilden
- ◆ Zwei Lösungswege
 1. Erfinder mit Autoren von wissenschaftlichen Artikeln abgleichen und darüber Verbindung zu Innovationsefeldern herstellen (z.B. Klassifikationsterme der Artikel)
 2. Patente basierend auf ihren Inhalten gruppieren und dann Innovationsefeldern zuordnen

- # Problem 1: Autor-Identifizierung
- ◆ Nur Namen der Autoren bei Publikationen und Patenten verfügbar (meist keine ORCIDs)
 - ◆ Mehrstufiger Clustering- und Matching-Prozess basierend auf Namen, Affiliation, Patenttexte, Klassifikation, Ort, etc.

Lösungsansatz: Autor-Identifizierung



Problem 2: Topic Clustering

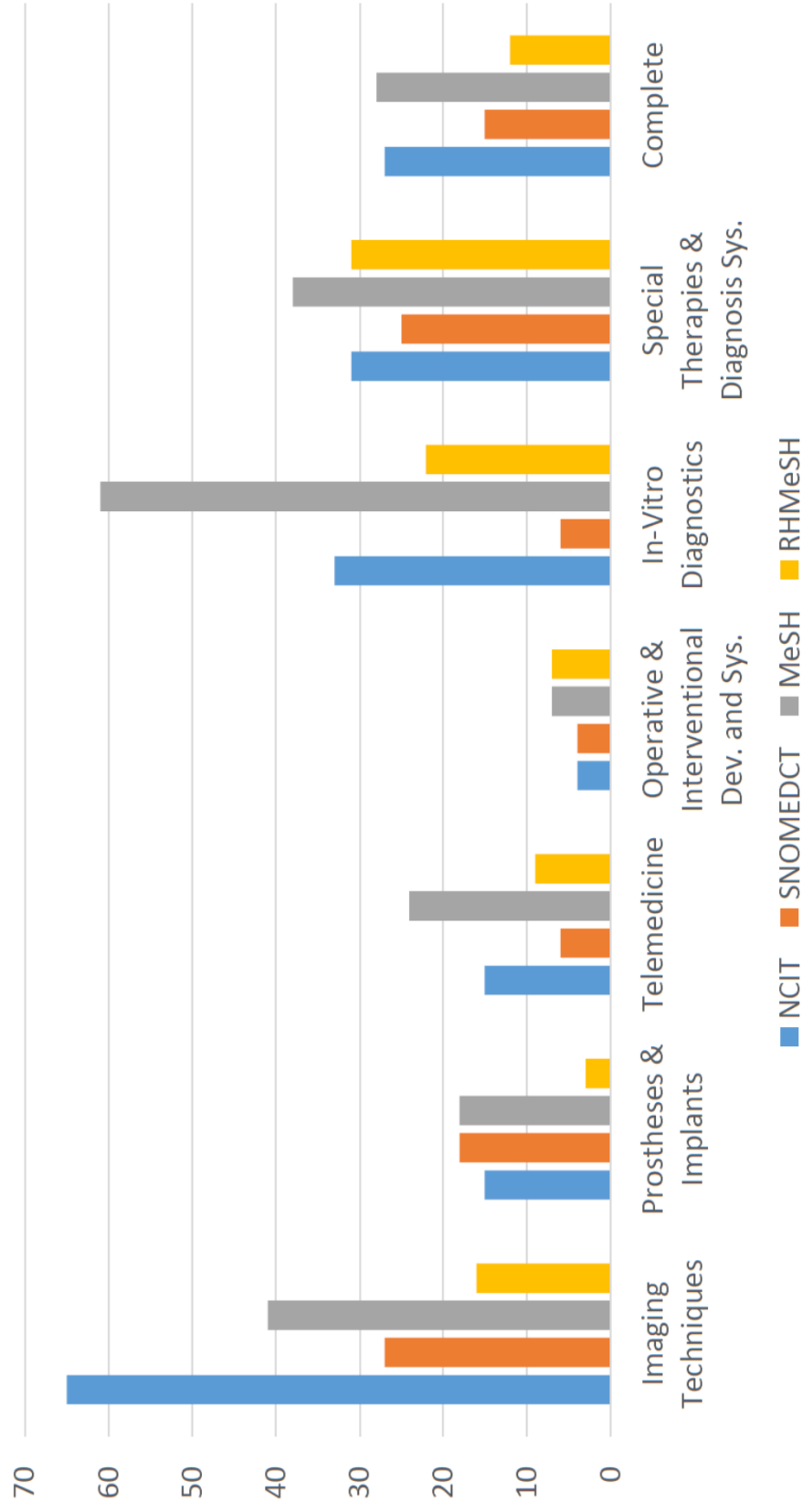
- ◆ Patente müssen anhand ihres Inhalts in Themengebiete gruppiert werden
- ◆ Häufigkeit von Begriffen (TF/IDF) in Abstract, Claims, Haupttext des Patents
- ◆ Lösung
 - Ähnlich wie für Autor-Identifizierung
 - Zuweisung von Begriffen zu Gruppen basierend auf Begriffshäufigkeit

Problem 3: Auswahl der Ontologien

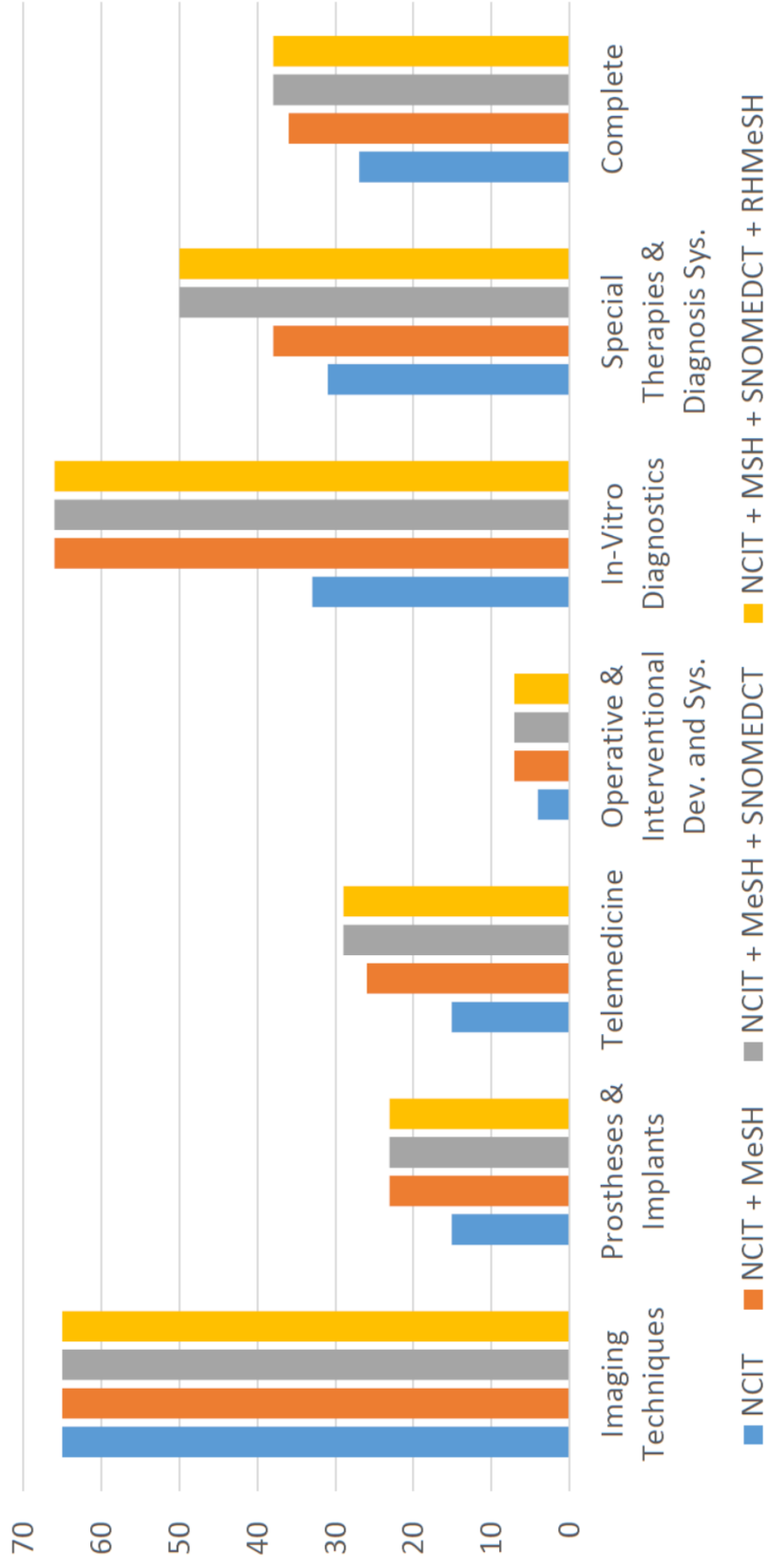
- ◆ Innovationfelder abstrakt, genaue strukturierte Beschreibung notwendig
- ◆ Publikationen und Patente schon klassifiziert (z.B. MeSH, IPC)
- ◆ IPC zu generisch

➔ Beschreibung der Innovationsfelder durch Verwendung von existierenden Ontologien

Abdeckung der Innovationsfelder in einzelnen Ontologien



Abdeckung der Innovationsfelder in mehreren Ontologien



Ontology Matching

- ◆ Abbilden der Begriffe aus den Publikationen bzw. Patentgruppen auf Ontologien
- ◆ Nutzung von klassischen Ontology-Matching-Werkzeugen (z.B. AgreementMaker, GeRoMeSuite) oder existierenden Mappings (z.B. Bioportal)
- ◆ Evaluierung durch Domänenexperten

Zusammenfassung

- ◆ Semi-automatische Methode zur Empfehlung von Projektpartnern
 - Patentanalyse
 - Clustering
 - Ontologie-Modellierung und -Matching
- ◆ Fokus auf Patente für produktorientierte Projekte in der Medizintechnik

Nächste Schritte

- ◆ Clustering und Autor-Identifikation wurden realisiert → Hohe Genauigkeit
- ◆ Ontologie für Innovationsfelder muss noch verfeinert werden
- ◆ Matching und Gesamtprozess umsetzen und evaluieren