

Aufbereitung und Analyse der Deutschen Forschungsberichte zur Indikatorik von Forschungsaktualität und -trends (TrenDTF)

Erstens versuchen wir mit *TrenDTF* die Erschließung des Korpus der Deutschen Forschungsberichte (DTF) durch Textmining-Verfahren (insbesondere Topic Modeling) zu verbessern. Dieser Korpus soll zu einer **Datengrundlage für neue und verbesserte Indikatoren für die Unterstützung der Evaluierung des Outcomes von Forschungs- und Entwicklungsaktivitäten** werden. Die Technische Informationsbibliothek (TIB) ist Depotbibliothek für Forschungsberichte verschiedener Ministerien und forschungsfördernder Institutionen, darunter das Bundesministerium für Bildung und Forschung (BMBF) und das Bundesministerium für Wirtschaft und Energie (BMWi). Stand Januar 2019 umfasst der TIB-Bestand rund 275.000 gedruckte und rund 80.500 elektronische Abschlussberichte.

Zweitens wird der DTF-Korpus um Komplementärdokumente wie Förderbekanntmachungen ergänzt, um eine vergleichende **Analyse der Absichten der Förderer und der daraus resultierenden Forschung** zu ermöglichen. Durch Abfrage verschiedener Förderer (BMBF, BMWi, DFG etc.) werden derartige Dokumente in einer Datenbank archiviert und erschlossen. Ferner werden Topic-Modeling-Verfahren auf die Ausschreibungstexte und Politik-Dokumente angewendet. Der Vergleich der so ermittelten Themen in diesen Komplementärdokumenten mit den Themen in den Forschungsberichten soll Aufschluss darüber geben, inwiefern die Forschungsprogramme zur Zielerreichung und zur Umsetzung der Missionen in der deutschen Wissenschafts- und Innovationspolitik beitragen. Es werden **Indikatoren zur regelmäßigen Analyse der Missionsorientierung der Förderprogramme** entwickelt und validiert. Darüber hinaus sollen Vergleiche der Themenprofile (Ergebnisse des Topic-Modeling) der Forschungsberichte und der deutschen Zeitschriftenveröffentlichungen in der bibliometrischen Datenbank „Scopus“ durchgeführt werden. Auf aggregierter Ebene kann auf dieser Basis abgeschätzt werden, wie sich die thematische Ausrichtung der Forschungsprojekte und des gesamten wissenschaftlichen Publikationsoutputs unterscheidet. Unterschiede in den Profilen können als Effekte der institutionellen Förderung bzw. anderer als Forschungsprojekte des Bundes interpretiert werden. Eine entsprechende Hypothese ist zu testen. Schließlich werden Artikel-Ansetzungen aus Wikipedia als Referenz-Vokabular zum Topic-Modeling erprobt, um auf dieser Grundlage auch die Aktualität, Abruf-Häufigkeit u.ä. frei zugänglicher Wikimedia-Daten zur Korrelation mit den Themen der DTF anwenden zu können.

Drittens wird durch Korrelation themenbezogener und bibliometrischer Daten aus DTF, Förderlinien und Wikipedia die Aktualität von Forschungsberichten im Zyklus der Aufkommens neuer Themen in der Forschungsliteratur, Förderpolitik sowie ggf. zu beobachtender Popularität der Themen anhand von Artikel-Entstehung und Aufrufzahlen in der Wikipedia-Datenbank beobachtet. Dieser Teil der Analyse zielt auf die Bewertung der Aktualität und der Dynamik der Forschungsberichte im Vergleich zu Wikipedia-Daten und zu Zeitschriftenveröffentlichungen ab. Es wird geprüft, inwiefern auf Basis der Forschungsberichte neue Werkzeuge für die **Detektion thematischer Trends** sowie Indikatoren zur Vermessung der **thematischen Aktualität** von Forschungsberichten sowie ihrer **thematischen Nähe zur Förderpolitik** entwickelt werden können.

Publikationen

Derzeit liegen noch keine Publikationen vor.

