

Wissens- und Technologietransfer an deutschen Hochschulen: Möglichkeiten KI-gestützter Auswertung von Informationen auf Universitätswebseiten*

Associate-Professor Dr. Torben Schubert

* Gefördert durch das BMFTR im Projekt KiWi (FKz.16WIT013B)

Schubert, T., Darold, D., & Kroll, H. (2025). Knowledge and technology transfer in German universities: An exploratory study leveraging information from university websites. *Socio-Economic Planning Sciences*, 102333..



Socio-Economic Planning Sciences

Volume 102, December 2025, 102333



Knowledge and technology transfer in German universities: An exploratory study leveraging information from university websites

Torben Schubert ^{a b}✉, Denilton Darold ^{a b}✉, Henning Kroll ^{a c}✉

Show more ▾

+ Add to Mendeley 🔗 Share 🗣 Cite

<https://doi.org/10.1016/j.seps.2025.102333> ↗

[Get rights and content](#) ↗

[Under a Creative Commons license](#) ↗

● [Open access](#)

Motivation

- Der Transfer von Wissen und Technologie aus der Wissenschaft in die Industrie und Gesellschaft ist von zentraler Bedeutung für nationale Innovationssysteme und das Wirtschaftswachstum (Edquist, 1997; Lundvall, 2007).
- KTT ist multidimensional (z. B. gemeinsame Forschung, Lizenzierung, Beratung, Personalaustausch) (Bozeman, 2000; Mowery & Rosenberg, 1999).
- Empirische Forschung stützt sich oft auf enge, leicht verfügbare Proxies wie Universitätspatente oder Co-Publikationen, wodurch viele Transferformen übersehen werden (Crespi et al., 2011; Jaffe, 1989).
- Viele relevante KTT-Interaktionen sind implizit oder undokumentiert, sodass es schwierig ist, groß angelegte Evidenz für die Vielfalt der Kanäle zu erhalten (Muhonen et al., 2020; Spaapen & van Drooge, 2011).
- Umfragen sind zwar hilfreich, aber in Umfang und Reichweite begrenzt, sodass grundlegende Fragen – Welche Arten von KTT gibt es? Welche werden von den Universitäten bevorzugt? – noch nicht ausreichend untersucht sind (Abreu et al., 2009; Benneworth et al., 2009).
- Universitätswebsites enthalten reichhaltige, unstrukturierte Informationen über KTT-Aktivitäten, die als Datenquelle bislang weitgehend ungenutzt geblieben sind (Medeiros & Souto, 2020).
- Motivation: Entwicklung umfassenderer, webbasierter Indikatoren zur Erfassung der gesamten „dritten Mission“, einschließlich unkonventioneller Transferformen, die über traditionelle, technologieorientierte Messgrößen hinausgehen.

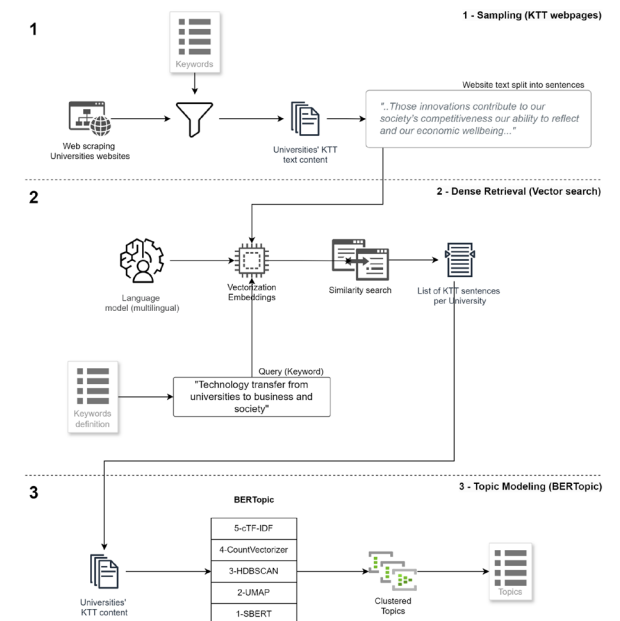


Kernidee

- Verwendung von Daten aus Universitätswebsites zur Identifizierung von KTT-bezogenen Mitteilungen auf den Universitätswebsites
- Verwendung von LLM-gestütztem Topic Modelling zur Extraktion relevanter Themen
- Abgleich mit anderen Daten auf Universitätsebene (Bibliometrie, Patentstatistiken, Registerdaten), um wichtige Fragen der KTT-Forschung erneut zu untersuchen
- Was prägt KTT? Das Umfeld? Die Universität? Die Präferenzen der Forscher?
- In welcher Beziehung stehen KTT und Grundlagenforschung zueinander

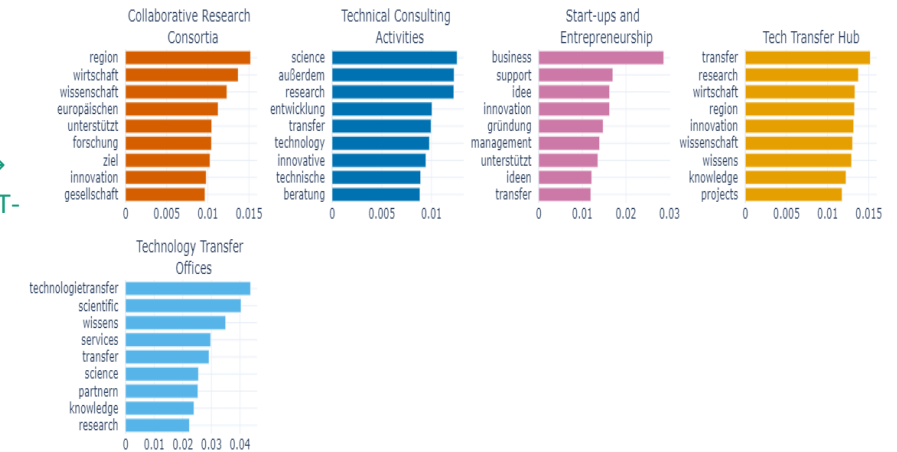
Methoden I

- Zielgruppe und Datenerhebung: Im Oktober 2023 wurden die gesamten Websites von 75 deutschen Universitäten (von insgesamt 90) mit Python-Scraping-Tools gescraped; die Inhalte wurden zur Einbettung in Satz-/Chunk-Einheiten (mind. 200 Zeichen) segmentiert.
- Zweistufige Relevanzauswahl: (1) keywordbasierte Seitenfilterung unter Verwendung eines iterativ entwickelten deutschen KTT-Lexikons; (2) semantisches Dense Retrieval mit Satz-Transformer-Einbettungen, um KTT-relevante Snippets abzurufen.
- Implementierung von Dense Retrieval: Verwendung des T-Systems-on-site/cross-en-de-roberta-sentence-transformer-Modells zur Einbettung von Text und erweiterten Keyword-Beschreibungen; FAISS-Ähnlichkeitssuche (L2-Distanz) mit einem Distanzschwellenwert von 12 zum Rangieren und Auswählen semantisch ähnlicher Sätze.
- Themenmodellierungspipeline: Anwendung von BERTopic (SBERT-Einbettungen → UMAP-Dimensionsreduktion → HDBSCAN-Clustering → c-TF-IDF für die Themenrepräsentation) auf den KTT-Korpus, um thematische Cluster zu erkennen.



Methoden II

- Modellierungsparameter und Auswahl: Vektorisierer-Mindestvorkommen = 30 (Bigramme zulässig), Themenumfang = 20 Unterseiten, HDBSCAN zur Verarbeitung von verrauschten, heterogenen Webinhalten ausgewählt; aus Gründen der Konsistenz wurde durchgehend dasselbe Satzmodell verwendet.
- Verfeinerung und Kennzeichnung der Themen: anfänglich 269 Themen → Auswahl der 48 häufigsten (>30 Vorkommen) → hierarchisches Clustering + manuelles Zusammenführen und Ausschließen → endgültiger Satz von 5 gekennzeichneten KTT-Dimensionen (Kooperationskonsortien, technische Beratung, Start-ups, Technologietransfer-Hub, TTOs).
- KTT-Indikatoren auf Universitätsebene: Für jede Universität wurden die BERTopic-Themenwahrscheinlichkeiten über alle Seiten gemittelt, jede Themenbewertung anhand des themenweiten Mittelwerts normalisiert, die normalisierten Bewertungen der fünf Themen zu einem zusammengesetzten Index aggregiert und eine logarithmische Transformation angewendet.
- Abgeglicherer Datensatz: Zusammenführung der aus dem Internet abgeleiteten KTT-Indikatoren mit demografischen Daten der Universitäten (Statistisches Bundesamt), bibliometrischen Daten (Scopus), Patentdaten (PATSTAT) und regionalen Wirtschaftsvariablen (INKAR); endgültige Analyseprobe = 60 Universitäten für Regressionsanalysen.
- Validierungs- und Analysestrategie: Bewertung der Validität anhand von Korrelationen mit traditionellen KTT-Maßnahmen (Patentierung, gemeinsame Veröffentlichungen von Universitäten und Industrie); Verwendung deskriptiver Statistiken und explorativer Regressionen (Halb-Logarithmus für KTT als DV; Log-Log für Forschungsergebnisse) unter Berücksichtigung von Forschungs-, Organisations- und regionalen Kovariaten.



Beispiele

- **Forschungskonsortien**

- „...ein Konsortium vereint Universitäten und Forschungseinrichtungen, um Lösungen für intelligente Produkte und Produktionsprozesse zu entwickeln.“ „...regionale Transformation beinhaltet die Integration lokaler Industrien in forschungsorientierte Innovationen, um nachhaltige Geschäftsmodelle und Prozesse zu schaffen.“

- **Technische Beratung**

- „...bietet eine Reihe von Beratungs- und Forschungsdienstleistungen, in der Regel für externe Partner oder in Zusammenarbeit mit diesen, um praktische Probleme zu lösen.“ „Das Team bietet externe Beratung und Fachwissen zur Optimierung industrieller Prozesse, insbesondere in den Bereichen Maschinenbau und Fertigungstechnik.“

Beispiele II

- Start-ups und Unternehmertum

- „... eine praktische Gelegenheit für Studierende, die Realisierbarkeit ihrer Geschäftsideen zu testen und eine Denkweise für die Entscheidungsfindung in risikoreichen Start-up-Umgebungen zu erwerben.“ „... das Ergebnis des viertägigen Coachings ist ein individuell zugeschnittener Start-up-Fahrplan mit Empfehlungen für weitere Schritte zur Gründung eines Unternehmens.“

- Tech Transfer Hub

- „...die Abteilung für Wissens- und Technologietransfer (WTT) fungiert als Brücke zwischen Wissenschaft und Industrie und fördert die Zusammenarbeit zwischen Universitätsmitgliedern und externen Partnern.“ „...bietet umfassende Unterstützung und Beratungsdienste zur Förderung von Innovationen und zum Transfer von Forschungsergebnissen.“

- Tech Transfer Offices

- „...die Abteilung für Wissens- und Technologietransfer (WTT) fungiert als Brücke zwischen Wissenschaft und Industrie und fördert die Zusammenarbeit zwischen Universitätsangehörigen und externen Partnern.“ „...bietet umfassende Unterstützung und Beratungsdienste zur Förderung von Innovationen und zum Transfer von Forschungsergebnissen.“

Ergebnisse I

Variable	Obs	Mean	Std. Dev.	Min	Max
KTT coop. research consortia	75	1.00	1.00	0.00	6.18
KTT techn. consulting	75	1.00	0.53	0.00	2.41
KTT start-ups	75	1.00	0.58	0.00	2.09
KTT regional techn. transfer	75	1.00	0.54	0.00	2.70
KTT TTOs	75	1.00	1.02	0.00	8.59
KTT topic diversity	75	1.00	0.20	0.00	1.11
KTT web-pages	75	147.11	340.04	1.00	2664.00
10%-excellence rate	70	0.16	0.04	0.04	0.32
Citation rate	74	1.82	0.77	0.00	5.20
Share industry copublications in %	74	12.56	11.98	0.00	100.00
Publication intensity	64	0.53	0.42	0.00	3.01
University patenting intensity	64	0.18	0.18	0.00	0.60
Personnel expenses	65	282725.00	294353.00	0.00	1100000.00
Share of engineering graduates in %	65	20.33	24.43	0.00	77.26
Population density	75	1454.07	1250.48	42.52	4777.04
GDP per capita in 1000 Eur.	75	49.75	20.29	22.80	97.58
High-tech employment (absolute)	75	75.02	55.32	17.70	225.20



Fig. 3. Geographic distribution of topics' prevalence

Ergebnisse II

		1	2	3	4	5	6	7	8
1	KTT coop. research consortia								
2	KTT techn. Consulting	0.24							
3	KTT start-ups	0.40	0.74						
4	KTT regional techn. Transfer	-0.02	0.35	0.57					
5	KTT TTOs	-0.26	-0.34	-0.31	-0.16				
6	KTT topic diversity	-0.19	0.11	0.26	0.38	0.23			
7	KTT web-pages	-0.17	-0.11	-0.20	-0.25	-0.29	0.13		
8	University patenting intensity	0.04	0.22	0.19	0.07	0.01	0.22	0.20	
9	Share industry copublications in %	0.39	0.07	0.14	-0.08	0.21	0.25	0.08	0.50

Ergebnisse III

- Positiver Zusammenhang zwischen KTT und bibliometrischer Performanz (Quantität und Impact)
- Zusätzlicher Effekt bei kooperativen Forschungskonsortia

	Log 10%- excellence rate	Log citation Rate	Log publication intensity
Log KTT coop. research consortia	0.40347*** (4.45)	0.29856*** (4.00)	0.56857* (1.78)
Log KTT techn. consulting	-0.03549 (-0.25)	0.04572 (0.40)	0.24564 (0.50)
Log KTT start-ups	-0.29161* (-1.75)	-0.10578 (-0.77)	-0.07370 (-0.13)
Log KTT regional techn. transfer	0.12662 (0.88)	0.09039 (0.77)	0.13162 (0.26)
Log KTT TTOs	0.05958 (0.45)	0.14342 (1.32)	0.15381 (0.33)
Log KTT topic diversity	1.43248*** (4.04)	-0.07408 (-0.25)	-1.29526 (-1.04)
KTT web-pages	0.00020** (2.09)	0.00022*** (2.75)	0.00025 (0.73)
Population density	-0.00004 (-1.38)	-0.00002 (-0.71)	0.00011 (1.05)
GDP per capita in 1000 Eur.	-0.00014 (-0.10)	-0.00006 (-0.06)	-0.00023 (-0.05)
High-tech employment (absolute)	0.00069 (1.04)	0.00005 (0.10)	-0.00465** (-2.01)
Personnel expenses	0.00000 (1.18)	0.00000*** (2.90)	0.00000*** (3.53)
Share of engineering graduates in %	-0.00433*** (-3.89)	-0.00196** (-2.15)	0.00124 (0.32)
Constant	-3.00929*** (-15.55)	0.77137*** (4.85)	-0.73226 (-1.08)
Observations	60	60	60
R ²	0.650	0.460	0.350

Offene Fragen

- Was erfahren wir aus den Webseiten? Schönfärberei, konkrete Projekte oder strategische Absichten?
- Wie können wir das erkennen? Welche Art von Aktivitäten sehen wir? Hardcore-Technologietransfer oder die dritte Mission im weiteren Sinne?
- Was tun gegen das Risiko, dass Webseiten manipuliert werden, wenn solche Indikatoren häufiger werden?

Zusammenfassung I

- Wichtigste Erkenntnis: Es wurden fünf Themenbereiche für KTT-Aktivitäten an deutschen Universitäten identifiziert: Forschungsk Kooperationen, technische Beratung, Start-ups/Unternehmertum, Technologie-Transfer-Hubs und Technologietransferstellen.
- Konzeptioneller Fortschritt: Erweitert die Messung von KTT über Patente und gemeinsame Publikationen hinaus und stimmt mit früheren konzeptionellen/empirischen Arbeiten überein.
- Methodik: Datengestütztes Web-Scraping von Universitätswebsites zur ganzheitlichen Klassifizierung von KTT-Aktivitäten.
- Stärken: Deckt die gesamte Hochschulpopulation ab, ist kosteneffizient, schnell reproduzierbar und konzentriert sich auf Aktivitäten, die von den Institutionen veröffentlicht werden (wahrscheinlich strategisch oder erfolgreich). Einschränkungen: Webtexte sind heterogen – konkrete Projektbeschreibungen vermischen sich mit allgemeiner oder werblicher Sprache, was Herausforderungen bei der Klassifizierung und den Informationsinhalten mit sich bringt; einige regionale Korrelationen waren schwach.

Zusammenfassung II

- Nächste Schritte: Bereinigung der Korpora, um allgemeine Textsnippets zu entfernen, Nutzung von LLMs zur Unterscheidung zwischen werblichen und inhaltlichen Inhalten und Durchführung qualitativer Studien zur Kommunikation von Universitäten.
- Forschungspotenzial: ermöglicht feinere und wiederholbare (Panel-)KTT-Indikatoren, unterstützt Analysen von Synergien nach KTT-Typ und eröffnet die Möglichkeit für kausale/Interventionsdesigns (z. B. Differenz-in-Differenzen).
- Gesamtimplikation: Webbasierte Indikatoren können die quantitative KTT-Forschung erheblich erweitern, indem sie frühere Datenlücken schließen, erfordern jedoch einen sorgfältigen Umgang mit textuellen Verzerrungen.

Vielen Dank für Ihre Aufmerksamkeit!

Kontakt

Associate-Professor Dr. Torben Schubert
Abteilung Innovations- und Wissensökonomie
Tel. +49 12 3456-357
torben.schubert@isi.fraunhofer.de

Fraunhofer-Institut für System- und Innovationsforschung ISI
Breslauer Straße 48
76139 Karlsruhe
www.isi.fraunhofer.de



Fraunhofer-Institut für System- und
Innovationsforschung ISI