Dec 11th, 12:00 AM

Investigating German Higher Education Institutions' Transfer Activities: New Measurements Based on Web Mining

Michelle Schmitt
University of Siegen, michelle.schmitt@uni-siegen.de

Christian Schröder
University of Siegen, christian3.schroeder@uni-siegen.de

Günter W. Beck
Siegen University, guenter.beck@uni-siegen.de

Arndt Werner
University of Siegen, arndt.werner@uni-siegen.de

ICIS 2023 Hyderabad, An Exploratory Study of German Higher Education Institutions Transfer Activities: New Measurements Based on Web Mining

Completed Research Paper

Michelle Schmitt
University of Siegen
Unteres Schloß 3; D-57072 Siegen
michelle.schmitt@uni-siegen.de

Christian Schröder
University of Siegen
Unteres Schloß 3; D-57072 Siegen
christian3.schroeder@uni-siegen.de

Günter W. Beck
University of Siegen
Unteres Schloß 3; D-57072 Siegen
guenter.beck@uni-siegen.de

Arndt Werner
University of Siegen
Unteres Schloß 3; D-57072 Siegen
arndt.werner@uni-siegen.de

Abstract

In recent years, higher education institutions (HEI) have expanded their involvement in diverse transfer activities (TA), extending beyond traditional teaching and research roles. These TA are often heterogeneous and informal, which makes measuring their full scope and effects challenging. In this article, we propose a new and straightforward to implement approach for mastering this task. In a first step, we theoretically derive three different dimensions of transfer, namely the transfer of knowledge, technology and personnel. For each of these categories, we develop an artificial intelligence (AI) optimized keyword list. Finally, we use these lists and apply web mining techniques and natural language processing (NLP) to measure TA from German HEI. To this end, we analyze a total of 299,229 texts from 376 German HEI websites. Our study shows that our proposed approach represents an effective and valuable tool for measuring TA from HEI and provides a foundation for further research.

Keywords: Academic engagement, higher education institutions, natural language processing, knowledge transfer, technology transfer, personnel transfer, web mining.

Introduction

Innovations are essential for sustainable development and economic growth – especially in knowledge-driven economies (Fagerberg and Srholec 2008; Ferreira and Carayannis 2019; Verspagen 2006). Higher education institutions (HEI) are considered as important catalysts, boosting a country's ability to innovate and thereby addressing societal issues (O'Shea et al. 2008; Wright 2014). To utilize their full transfer potential, initiatives stimulating academic engagement are required (Dolmans et al. 2022). Hence, HEI reframed their policies and governance in the last two decades (Acs et al. 2017; Hayter 2015), becoming more entrepreneurial and transfer orientated (Abreu and Grinevich 2013; Zhao et al. 2020). Equipping themselves with a set of institutions - such as technology transfer offices (TTO), entrepreneurship centers, or university-based incubators and accelerators - HEI aim on leveraging their full entrepreneurial potential (Rasmussen and Wright 2015), including the transfer of knowledge into industry and society (de Wit-de

Vries et al. 2019; Hayter et al. 2020). However, finding the right balance between their traditional missions of teaching and research and the growing interaction with the private sector remains a challenge.

In the literature, transfer is often understood as a codified and formalized process of commercializing innovative products and services stemming from HEI via licensing, patenting, or academic-spin offs (e.g., Azoulay et al. 2009; Lach and Schankermann 2004; Lockett et al. 2005; Phan and Siegel 2006), often referred to as technology transfer. Contrary to this, Dosi (1982) describes the transfer process as multifaceted, involving codification (formalization and documentation of knowledge), personalization (knowledge transfer through direct personal interactions), and hybridization (a combination of codification and personalization). Therefore, a broader understanding of transfer activities (TA) suggests that they can go beyond formal activities like patenting or licensing (D'Este and Patel 2007). Nowadays, a central component of transferring knowledge is the exchange and application of research results generated within academia for the private sector (Ward et al. 2009). Moreover, in line with the growing entrepreneurial orientation of HEI, their set of TA becomes more diverse, including for instance technical assistance, consulting activities, collaborative research, or transferring knowledge via conferences (Geuna and Muscio 2009; Karnani 2013; Schaeffer et al. 2020). These formal and informal transfer mechanisms are complemented by personnel transfer (going beyond the education of students), such as for example student participation in joint projects, internships and thesis or doctoral projects (Arvanitis et al. 2008).

Since TA have become so diverse and widespread (D'Este and Patel 2007), measuring them is more challenging than ever and mostly limited to codified forms (Perkmann et al. 2013). Therefore, new methods in research are needed to analyze how, where and to what extent knowledge spills over from HEI to other areas (Hayter et al. 2020). We propose that web mining provides a useful approach, especially for examining fragmented and non-codified data. While gathering text data and especially texts from websites gain attention in several fields of research like economics and business research (e.g., Guiso et al. 2015; Kinne and Axenbeck 2020; Obschonka et al. 2020; Prüfer and Prüfer 2020), their application in research concerning TA is rather new. A few studies leverage this technique with mixed results. Della Volpe and Esposito (2020: p. 235) for example analyze mission statements from Italian universities, finding that about 30% of them mentioned TA, mainly focusing on immediate economic outcomes. However, these mission statements may not offer a complete picture and could perpetuate a distorted view, as they often serve more as aspirational tools for public relations than accurate reflections of operational activities (Giesel 2007; Ulrich and Fluri 1995). Other studies focus, for instance, on the analysis of intellectual capital (Secundo et al. 2017), author-centered publications, grant or patent applications (Palmer 2010), or on a combination of publications and financial data (European Commission 2018). Recently, Klingler-Vidra and Chalmers (2023) utilized web mining to assess the influence of university alumni in policymaking roles, revealing that their education impacts startup-centric policies. These studies, however, primarily address the traditional teaching and research missions of HEI. To date, a comprehensive analysis focusing on the third mission of TA in HEI is still missing.

While there is a growing tendency towards the institutionalization of TA, assessing a HEI third mission performance remains challenging. Existing rankings, such as the Times Higher Education World Ranking or the Academic Ranking of World Universities focus on teaching and research (Marhl and Pausits 2013), excluding third mission activities. However, measuring TA is a highly important task driven by policymakers trying to foster the role of universities as key actors of economic and cultural growth (Secundo et al. 2017). Hence, our study aims on increasing the understanding about transfer mechanisms from HEI into the global arena and vice versa, following the call for a holistic measurement of TA (Etzkowitz 2016). Particularly informal and non-codified transfer forms are targeted, as they are frequently utilized in every-day life but often get neglected from studies due to their lack of measurement (Barbini et al. 2021).

In this paper, we develop an approach that uses web mining and natural language processing (NLP) to measure communicated TA by HEI and apply it to German HEI. In a first step, we use ARGUS (a tool for automated large-scale web-scraping) to scrape 376 HEI websites and up to more than 2000 webpages assigned to each university website. This enables us to create a rich dataset of 299,229 texts. We then utilize reviewed literature, AI-optimization and manual quality controls to develop a keyword list categorizing three different types of TA: transfer of knowledge, technology and personnel. Based on this, we match the text corpus with our keyword list and analyze how frequently German HEI communicate TA within the respective categories. Moreover, we validate the applied keyword lists by measuring the dimensionality (i.e., independence) of our three proposed categories. The keyword list is open source and provided via the

following GitHub link (https://github.com/mischmitt/TAGHEI) to the public. Our study provides researchers with a practicable and usable dictionary and hopefully inspires future integration of web mining in transfer research. Finally, we apply our new measure by analyzing if and how HEI ownership and HEI research orientation is related to the frequency of HEI TA.

Our study offers a holistic approach for the measurement of knowledge, technology, and personnel transfer using data science methods, providing initial evidence that this procedure can serve as a foundation for future research. We demonstrate that data science and natural language processing techniques are well-suited for extracting non-codified information in the context of HEI TA. Combining web mining and NLP allows us to capture a comprehensive view of TA from German HEI, quantifying activities often neglected in the literature. Consequently, our study addresses the call for incorporating informal TA into academic entrepreneurship (AE) research (Abreu and Grinevich 2013; Barbini et al. 2021; D'Este and Patel 2007; Hayter et al. 2020, Perkmann et al. 2013) using state-of-the-art data science methods. Contrary to previous web mining studies in the field (e.g., Della Volpe and Esposito 2020; Klingler-Vidra and Chalmers 2023), this is the first study highlighting the role of formal and informal third mission activities in HEI through web mining. In summary, our study makes three interconnected contributions: First, we develop an open-access dictionary for the analysis of communicated TA in the context of German HEI by employing a distinctive combination of both inductive and deductive methods. This involves computational work intertwined with human input. The reproducible methodology encourages replication in other regions. Second, while existing measures primarily focus on the first and second missions of HEI—teaching and research—our study introduces a web-mining-based metric for assessing the third mission, aiding policymakers in evaluating HEI impact. Finally, we highlight the importance of non-codified TA, advocating for a multi-dimensional evaluation that encompasses knowledge, technology, and personnel transfer. Our results align with prior studies, confirming web mining's utility for extracting non-numeric data.

Related literature

The three academic revolutions of HEI

HEI underwent three academic revolutions in the past decades, resulting in drastic changes regarding their roles (for an extensive overview see Compagnucci and Spigarelli 2020). While the first academic revolution (Etzkowitz and Webster 1998) combined research and teaching (Urdari et al. 2017), the second academic revolution (Etzkowitz 1998, 2003) focused on commercialized and practically used knowledge creation in HEI (Rolfo and Finardi 2014). Consequently, universities introduced a large set of entrepreneurial support programs and activities. These range from patenting and licensing activities, the promotion of academic spin-offs to investing equity into start-ups (Mariani et al. 2018; Mowery et al. 2004; Siegel 2006). Thus, HEI contribute proactively to the social and economic development (O'Carroll et al. 2006) and interact with the society beyond their boundaries (Molas-Gallart et al. 2002). These developments opened the path for the third academic revolution. Consequently, HEI are nowadays key contributors to cultural and economic growth acting in a global environment (Svensson et al. 2012), including cooperation with the industry and society at large (Etzkowitz and Leydesdorff 2000; Vorley and Nelles 2008). Today, HEI can be characterized as institutions engaged in industry and society, strongly embedded in the global arena (Svensson et al. 2012) acting far beyond their traditional mission of teaching and research. Because universities now need to legitimize their activities not just to academia, but also to external stakeholders (Berman 2012; Fini et al. 2010; Murray 2010), they diversify funding, form research partnerships, and facilitate personnel exchanges with the industry (Abreu and Grinevich 2013; Baldini et al. 2015; Lockett et al. 2005; Siegel et al. 2007). This diversification makes measuring multifaceted TA crucial for aligning with stakeholder demands. In the case of Germany, state ministries outline TA-focused strategic goals in university laws, underscoring the growing policy emphasis on knowledge and technology transfer, as well as social engagement. While existing ranking systems focus mainly on teaching and research, both universities and policymakers are keen to find comprehensive methods for measuring their third mission activities (Montesinos et al. 2008). Research indicates that HEI still lack detailed data and tools required to evaluate their success in third mission activities (e.g., Wright et al. 2004), resulting in a call for a new measurement of third mission activities (Etzkowitz 2016).

Categorizing TA

TA are a multifaceted phenomenon. To reduce complexity, the activities can be distinguished by the underlying subject that is transferred between academia and the private sector. Based on the literature, three categories emerge. First, knowledge transfer encompasses know-how, theoretical knowledge, and procedures (Dosi 1982). Second, technology transfer involves research results and technological knowledge, emphasizing application and commercialization with codified and quantifiable characteristics (Azoulay et al. 2009; Lach and Schankermann 2004; Lockett et al. 2005; Phan and Siegel 2006). Lastly, personnel transfer centers on personal networks, experiences, and expertise exchange through (international) mobility (Aman 2020; Inkpen and Tsang 2005). A detailed overview of these three categories follows.

Knowledge transfer is defined as the transmission of expertise, experiences, and other forms of non-codified activities between actors, such as academic institutions and the private sector (Dosi 1982). Davenport and Prusak (1998) describe knowledge as a fluid mix of framed experience, values, contextual information, and expert insight, providing a framework for evaluating and incorporating new information. This “fluid mix” encompasses a wide range of assets, necessitating the further specification and contextualization of knowledge transfer. Dosi (1982) distinguishes between practical and theoretical knowledge application. Practical knowledge aims to solve existing problems, while theoretical knowledge generates new insights without immediate application. This distinction is consistent with the broad transfer perspective, which includes activities such as joint publications with industry scientists, industrial consulting (Link et al. 2007), data sharing, collaborative R&D projects (Morandi 2013; Schultz et al. 2021), and educational services (Bektaş and Tayauova 2015; Giones et al. 2022). These activities can differ in applicability and research scope, ranging from basic to applied research. However, they share the common feature of being based on new, informal, and hard-to-quantify knowledge. Institutional offerings that facilitate knowledge transfer include incubators (Blank 2021), university entrepreneurship programs (Menzies 2000), and accelerators (Crişan et al. 2021).

Technology transfer contains all activities that focus on the transfer of a certain technology. It consists of codified and numeric activities. Technology transfer is therefore closely intertwined with the narrow transfer definition. While the term is frequently discussed in AE literature, we follow the idea of technology as an “application of knowledge, tools, and skills to solve practical problems and extend human abilities” (Johnson 1989: p.1). In line with the narrower transfer understanding, transfer of technology aims on solving explicitly formulated practical problems. The most prominent examples of this consist of licensing, patenting or royalty agreements (e.g., Azoulay et al. 2009; Lach and Schankermann 2004; Phan and Siegel 2006; Thursby et al. 2001), research joint ventures (Hess and Siegwart 2013), university-backed start-ups or academic spin-offs (Lockett et al. 2005; Meoli et al. 2019). Technology transfer is contractually governed and supported through TTOs (Faccin et al. 2022) or science and technology parks (Albahari et al. 2022).

Personnel transfer consists of personnel mobility between academia and the private sector. Personnel transfer does not denote the absolute number of alumni from HEI to the private sector, as that aligns with the HEI primary teaching mission (e.g., Klingler-Vidra and Chalmers 2023) - instead, it signifies practices of knowledge exchange involving university staff, students, and individuals outside academia. Often referred to as "transfer over heads", its impact remains largely unexplored due to its non-codified nature. Empirical evidence from Italy suggests social proximity influences student entrepreneurship (Barbini et al. 2021) and Ph.D. student engagement in entrepreneurship (Muscio et al. 2022), but the frequency and occurrence of personnel exchange between universities and the private sector are still underexplored. We aim to identify these mechanisms by measuring personnel transfer as an independent category. Examples of personnel mobility include industry employee training, postgraduate industry training (Perkmann and Walsh 2007), internships (Lee 2000), joint thesis projects, and personnel exchange through consulting (Arvanitis et al. 2008) or guest lectures (Bischoff et al. 2018). Recognizing that personnel mobility is multidirectional, it is essential to consider the broader scope, encompassing university knowledge flow into hybrid-entrepreneurship and intrapreneurship, leveraging incumbent firms' competencies (Alsos et al. 2023). Knowledge spillovers from the private sector to academia also occur, with university heads increasingly applying business-related knowledge in academic contexts. As entrepreneurial universities grow, academics are expected to commercialize research (Leitner et al. 2021). Given the well-integrated networking events between HEI and private sector actors (Backer et al. 1980), personnel transfer is a central component of TA that warrants thorough analysis.

Antecedents of TA

German HEI are heterogeneous in terms of size, age, ownership, location and research focus. Depending on their research focus, regular universities and universities of applied sciences have distinct characteristics. In eight of the sixteen German states universities of applied science are not entitled to confer doctorates and many professorships are not equipped with doctoral or post-doctoral students, affecting the transfer output in these institutions. In our analysis, we distinguish regular universities, universities of applied sciences, and universities for arts and music. Also, HEI can be publicly funded and operated, privately funded and managed, or church-owned. This diversity impacts the type of research conducted and the resources provided to researchers. Size varies significantly among German HEI, with some accommodating over 50,000 students (e.g., LMU München, Universität zu Köln (universities); IU Internationale Hochschule, Hochschule für Ökonomie & Management (universities of applied sciences)) and others having fewer than 20 students (e.g., Philosophisch-Theologische Hochschule Münster, Kolping Hochschule). Lastly, German universities exhibit a wide range in their history and tradition, with the oldest university, Ruprecht-Karls-Universität Heidelberg, founded in 1386, and the youngest established in 2021.

Regarding research orientation, universities of applied sciences are expected to be more engaged in regional transfer compared to regular universities (Arvanitis et al. 2008; Jaeger and Kopper 2014). Regular universities' basic research in physical science and engineering leads to more formalized TA (Abreu and Grinevich 2013; O'Shea et al. 2005). In contrast, TA from universities of applied sciences and in arts, humanities, and social sciences tends to be less formalized, sometimes not even considered entrepreneurial or transfer-related (Abreu and Grinevich 2013). When comparing generalist and specialist universities, generalist institutions seem to have better conditions for commercialization due to higher internal knowledge spillover and interdisciplinary collaboration opportunities (Bonaccorsi et al. 2022; Giuri et al. 2019). Regarding HEI ownership, public universities primarily depend on state funding (Berbegal-Mirabent et al. 2015), while private universities, with greater institutional autonomy (Fini et al. 2017), capitalize on student numbers and other revenue sources. Due to their commercial orientation, private HEI often maintain closer and more frequent ties to industry partners (Belenzon and Schankerman 2009). Thus, they appear more active in TA compared to public HEI (Siegel et al. 2003), which prioritize basic research and teaching over research and TA (de la Torre et al. 2017). This approach contrasts with the marketability of private institutions (Vinig and Van Rijsbergen 2009). However, the debate continues, as some studies report greater success in marketable transfer outputs for state universities (Bonaccorsi et al. 2022).

Development of a new framework for TA research

Creation of a new dataset for TA from German HEI

While previous related studies focus on the analysis of mission statements or other parts of TA communication, we aim to include full websites of HEI. For this purpose, we compile a new dataset of texts from German HEI websites. Based on a list from "Hochschulkompass" (a study guide for potential students), we extracted basic information of 423 German HEI. The information on this website is authorized and updated by the HEI and can be accessed using this link: https://www.hochschulkompass.de/en/higher-education-institutions/downloads.html. The list includes the name of the HEI, the URL of its website, their research orientation and ownership structure as well as the state in which the HEI is located. Cross-checking the list from "Hochschulkompass" with manual web search and official statistics, we can assume the population of HEI in Germany to be n = 432. Out of these we were able to access and scrape the websites of 376 HEI in sufficient quality. More specifically, this means that we could scrape websites with at least ten non-empty and human readable texts as websites with less texts are likely misrepresented. This represents 87.04% of all state-approved HEI in Germany, where students can get an officially recognized degree. This mostly includes German HEI except for the Touro College Berlin, which is a German branch of a US HEI-network.

Collecting texts from each website was done in an automated manner by web-scraping using the web-scraping tool ARGUS (Automated Robot for Generic Universal Scraping) developed by Kinne and Axenbeck (2020). This tool is specifically designed for efficient large-scale web-scraping and enables us to extract numerous texts from the visited pages thereby covering full or nearly full websites. ARGUS takes a list of URLs as input and crawls the website automatically starting with the first page that it is given until it has

gathered all available pages on the website or until it reaches a given scraping limit, which commands the program to stop the procedure when reaching a defined number of pages. Crawling the links is not randomized, but follows certain heuristics set by the user, e.g., prioritize following short links or links that indicate contents in a certain language. We chose to prioritize links that indicate German contents. However, this does not limit ARGUS in general or our search to only one language. While ARGUS can scrape any textual content in any language, our search results in a dataset with mostly German, but also English texts. At the end of the scraping procedure, ARGUS generates a CSV file as output that contains the URL of the requested webpage, the error code for each webpage and the text of the webpage in case of successful scraping (without errors). This is the information in the output that we use for our study. For more details concerning parameters and output variables in ARGUS output, see the detailed description of the program in the GitHub repository (available here: https://github.com/datawizard1337/ARGUS).

As input, we use the URLs of the homepages of HEI that we gather through "Hochschulkompass" and manual web search. To avoid malfunctioning and manual interventions into the process, we set a scraping limit of 1000 pages. This proved to be a reasonable number to prevent infinite running time and at the same time scrape full or nearly full websites. After the first scraping procedure, we then again scrape those pages that respond with error codes. In this way, we gather texts from some of the pages that could not be reached in the first round. Once again, the scraping limit is set to 1000 pages, so ARGUS may crawl and request parts of the website that haven't been visited before. This step is repeated with remaining pages and a scraping limit of 1. Consequently, we are able to extract more than 2000 texts for some HEI.

Finally, we merge all output files and delete duplicates and pages with empty text. The resulting dataset consists of 299,229 raw texts from 376 HEI. Our dataset combines the basic information on names, postcodes, states, research orientation and ownership structure with the raw texts generated by the ARGUS web scraping procedure. The raw texts therein are only cleaned of HTML formatting and all capital letters are transformed to lowercase letters. These two elementary cleaning steps are already applied during the merging process of all ARGUS output files. Although the scraping procedure does not result in high quality output for all HEI, we can still cover 87.04% of the population of state-approved HEI in Germany and build a highly representative dataset for analyzing TA in Germany.

Development of a new measurement for TA

The main purpose of this study is to develop a new measurement for TA that benefits greatly from Big Data and text mining and hence captures so far neglected or underrepresented information. For our first exploratory approach, we follow the method of computer-aided text analysis (CATA), a procedure formalized by Short et al. (2010). The general idea behind this approach is that each of the keywords in a specified list is connected to the TA field. Consequently, a higher number of observed keywords can be used as a proxy for a higher commitment of the HEI towards TA. In line with this approach, we first set up a list of keywords (dictionary) and then choose a suitable methodology for the measurement of references.

Construction of the TA dictionary

Following Short et al. (2010), we use a procedure of deductive and inductive steps to construct a dictionary of TA-related keywords that allows for an analysis of the three TA dimensions. In a first deductive step, we construct a list of keywords associated with the broad topic of TA. Based on terms from related literature, manually picked HEI websites and expertise, each author suggests suitable terms and revises the suggestions of the other authors. Finally, we agree on a list that does not only contain single words, like "Entrepreneurship", but also n-grams consisting of up to four words. We include German as well as English terms, accounting for the two most common languages in German higher education. Moreover, we categorize the keywords according to the described theoretical framework in the prior sections into terms related to the categories: transfer of knowledge (e.g., knowledge transfer, entrepreneurship center), technology (e.g., technology transfer, patent) and personnel (e.g., transfer of personnel, internship).

After the first manual step of construction, we widen the list inductively using the information from the gathered texts. Using AI, we aim to ensure that our lists also reflect terms from the English and German language that are related to TA that have so far been neglected. Therefore, we program an algorithm to search for terms in an exhaustive list of German and English vocabulary that are close to our chosen keywords in terms of similar spelling. This vocabulary is created based on the gathered website texts as well

as the NLTK Wordnet corpus (Bird et al. 2009) and a collection of English words on GitHub (available online https://github.com/dwyl/english-words) for English terms and based on a selection of corpora from Wortschatz Leipzig (Goldhahn et al. 2012) for German terms. The corpora of Wortschatz Leipzig can be downloaded from this repository: https://wortschatz.uni-leipzig.de/de/download/German. From there we download the corpora mixed-typical 2011 1M, news 2021 1M, newscrawl 2018 1M, newscrawl-public 2019 1M, web 2011 1M and web-public 2019 1M to gather a broad vocabulary from different fields of language usage. In this way we consider a vast vocabulary, which enriches our keyword list and does not limit its use specifically to the HEI websites that are collected during this study.

The algorithm for the list enhancement can be described as follows. For each of the entries in the keyword list the algorithm first searches for terms that contain the keyword at hand, e.g., it adds “Entrepreneurship” if the keyword is “Entrepreneur”. Then it searches for all similar keywords and phrases in the collected texts and adds them to the list. Hereby, we rely on a similarity score calculated using the token set ratio function of the Python library “FuzzyWuzzy” (see documentation https://pypi.org/project/fuzzywuzzy/). More explicitly, the algorithm computes a similarity score based on Levenshtein distance between each term in the preliminary keyword list and each term in the vocabulary list. It then appends all terms from the vocabulary with a similarity of at least 80% (compared to the keyword at hand) to the preliminary keyword list. In this way the algorithm can add for example the term "innovationtransfer" (if it is not part of the list) which is similar to "informationtransfer". As the list of keywords is split into three categories, the algorithm assigns the newly detected terms to the corresponding category of the keyword at hand. If for example “licensing” is added as a similar word to “license” it is assigned to the same category (technology transfer).

This procedure has the advantage that we do not miss information due to errors in stemming, i.e., comparing only the stems of the words. Especially in the German language the algorithm might stem words incorrectly or not at all due to unknown or uncommon endings. On the other hand, this algorithm might add irrelevant keywords to the list, like “patient” as a close match for “patent” or assign them to the wrong category. Therefore, the resulting list is checked manually and falsely added keywords are deleted or assigned to another category. Again, the authors revise the suggestions for removal or re-assignment of the other authors and agree on a final list. The similarity parameter of 80% proves to be the most efficient leaving a sufficiently large set of tokens for the authors to choose from while at the same time not overloading the list and making it impossible to handle manually.

As a result of this procedure, we generate a dictionary that contains keywords based on expertise, literature and website searches, but is additionally enriched by relevant terms from an exhaustive vocabulary of the English and German language. In a last step, all keywords are transformed in the same way as the terms in the text. The exact procedure is described in the section “Text Preprocessing”. The core of our deductively created list consists of 182 terms, of which 68 refer to transfer of knowledge, 82 to technology transfer and 33 to transfer of personnel. Modifying the list using the described procedure results in a list of 1381 keywords, of which 369 belong to the category transfer of knowledge, 793 to technology transfer and 231 to transfer of personnel. The complete dictionary can be found in the GitHub repository of this project (https://github.com/mischmitt/TAGHEI). As the numbers suggest, there are overlaps of keywords between the different categories. Whenever a keyword can be used in more than one category, we include it in both lists instead of assigning it to only one category. This is due to the complex nature of languages and the fact that one keyword might describe an activity focused on transferring knowledge as well as one that involves mobility of personnel. Nevertheless, we can validate the dimensionality of our theoretical construct and our dictionary as described in the section “Validity of the proposed research framework”.

Choice of Methodology

Our objective is to proxy TA by measuring the frequency with which HEI communicate such topics on their websites. A straightforward approach to achieve this goal is to simply count the occurrence of given terms related to TA. In NLP literature this measurement is known as term frequency (TF). By applying this idea, we assume that all references are used in a positive context and that communication is positively related to activity. This assumption builds on the ongoing debate on the positive link between thoughts and language (communication) (e.g., Boroditski et al. 2011; Lakoff and Johnson 2008;) that has its origin in the Sapir-Whorf-Hypothesis (Sapir 1944; Whorf 1956). Although this intuitive measure of TF is widely used it can also vary significantly with the length of texts. NLP literature provides various ways to address this issue and assign weights to words. Apart from simple TF, one of the most popular measures in NLP is the term

frequency inverse document frequency (TF-IDF) or variations based on it, such as measures proposed by Baena-Garcia et al. (2011), Beel et al. (2017) and Bun and Ishikuna (2001). The general idea behind TF-IDF is to account for the fact that words that appear frequently in many documents do not contain much additional informational value and therefore receive a lower weight (Robertson 2004). While this is useful to measure how well a specific term characterizes a text and can therefore be seen as a measure of representation (Aizawa 2003), our analysis requires a measure of popularity such as TF. We therefore choose to calculate TF as proxy for TA and commitment. To introduce comparability across different lengths of texts it is a common practice to measure the TF per 100 or 1000 tokens in the text (Bodnaruk et al. 2015; Kindermann et al. 2021; McKenny et al. 2018; Moss et al. 2018; Vaupel et al. 2022). For a better readability, we decide to measure the TF per 1000 tokens. In line with this approach, we calculate the TF according to the following equation:

$$TF_{t,d} = tf_{t,d}/n_d*1000$$

where t is a given term (word or n-gram), d is a document, $tf_{t,d}$ is the TF of term t in document d and n_d is the number of all terms used in the respective document. Consequently, $TF_{t,d}$ is the TF per 1000 tokens for term t in document d. In this way we provide comprehensive results that are comparable across different lengths of texts.

Calculation of the TA measurement

Text Preprocessing

Before calculating TF, data cleaning and preprocessing is necessary to guarantee reliable and unbiased results. We follow the standard cleaning procedures in NLP, including tokenizing, stemming and removing punctuation and digits, to bring the texts into an adequate format for analysis. Using the large vocabulary of English and German words created earlier, we delete terms that are probably non-existent in real languages, such as leftovers from HTML formatting. In our preprocessing we moreover delete "stopwords" (English and German). Furthermore, we remove names and abbreviations of HEI and names of cities and locations as they are frequently present on HEI websites without contributing informational value for our analysis. As a result, each text is transformed to a list of terms with informational value. The same procedure is applied to the dictionary to ensure that the keywords therein match the tokens of the texts.

Calculation of TF and Relevance of TA

To get a first impression of the relevance of the broad topic of TA, we initially count all texts in general (i.e., in the whole corpus) and then on institution-level that contain at least one of the keywords specified in the full dictionary. For a better comparison we divide the resulting number by the number of all texts (in the corpus and at institution-level) and generate the measurement of the percentage of relevant texts. The calculation of TF is performed at the institution-level. For this procedure we combine all texts of a specific institution to one, regarding the complete website of an institution as one text. We then apply the formula for the TF to the texts. This is done for each term in the dictionary separately and in form of a „grouped" TF for the complete dictionary and its three categories. During this procedure any occurrence of one of the terms in the dictionary is counted per 1000 tokens. In this way we create TF measurements for all terms (complete dictionary) and terms referring to transfer of knowledge, technology and personnel.

Results on the TA of HEI in Germany

Descriptive Statistics

In this section we present our findings on TA in German higher education, starting with the most relevant descriptive statistics. Table 1 shows a summary of the dataset and the key results of TF and percentage of relevant texts calculated for the whole corpus. While almost half (45.28%) of the collected texts refer with at least one of the keywords to the overall topic of TA, all HEI in our sample do so in at least one of their texts. In other words, there is no HEI that does not refer to TA at all. Overall, almost 5 terms per 1000 in the corpus refer to TA, according to our specified dictionary. The most present category is transfer of knowledge with 2.19 terms per 1000, followed by transfer of technology with 1.6 terms per 1000, leaving transfer of personnel as the least present category with 1.48 terms per 1000.

Statistics	Value
Total number of texts	299229
Number of HEI	376
Average number of texts per institution	795.82
Number of relevant texts	135488
Percentage of relevant texts	45.28
Number of HEI without TA	0
TF for all terms	4.96
TF for transfer of knowledge	2.19
TF for transfer of technology	1.6
TF for transfer of personnel	1.48

Table 1. Summary statistics

Table 2 presents descriptive statistics focusing on the institution-level results. Regarding the TF of our dictionary and its three categories, the main finding from the summary statistics remains. The mean, median and maximum values show that transfer of knowledge is mentioned the most often of all three categories, followed by transfer of technology and then transfer of personnel.

Statistics	mean	sd	median	min	max
Number of texts	795.82	387.6	858.5	13	2224
Number of unique tokens on the website	22525.8	11384.2	16292.5	793	45171
Number of relevant texts	389.106	314.706	266	2	1414
Percentage of relevant texts	50.3006	28.5954	34.11	2.99	100
TF for all terms	4.96	4.51	3.98	0.28	38.23
TF for transfer of knowledge	2.19	3.1	1.31	0	29.33
TF for transfer of technology	1.6	1.85	1.09	0	18.74
TF for transfer of personnel	1.48	1.36	1.08	0	7.79

Table 2. Descriptive statistics based on institution-level calculations

Validity of the proposed research framework

When constructing dictionaries for CATA it is recommended to follow common procedures and strive to prove validity (Short et al. 2010). By building our multidimensional dictionary on theoretical foundations and combining the proposed deductive and inductive approach for setting up the list of keywords (Short et al. 2010), we follow the suggestions closely. Although we do not test for content validity by calculating rater's reliability scores from expert judgements (Short et al. 2010), we ensure that the authors agree on which words to keep in the list and in which category.

One way to validate dimensionality is to calculate correlations between the measures for the different dimensions (categories). Correlations between the measures of the different dimensions should be positive, but not excessively high (Short et al. 2010). Table 3 shows the correlation coefficients between the TF for each of the dimensions. As can be seen, the TF for each dimension is positively correlated to the TF of the other two dimensions, but with correlation coefficients below 0.5 this correlation is not strong. Visibly, the dimensions correlate more strongly to the overall topic of transfer as this comprises the sum of all keywords used in the study. In this way, we validate the dimensionality of our theoretical construct and our dictionary.

	TF for all terms	TF for transfer of knowledge	TF for technology transfer	TF for transfer of personnel
TF for all terms	1			
TF for transfer of knowledge	0.89	1		
TF for technology transfer	0,65	**0.42**	1	
TF for transfer of personnel	0.57	**0.30**	**0.17**	1

Table 3. Bivariate correlation coefficients between the TF of the different dimensions

Reflection of contextual factors of TA in the results

To demonstrate the meaningfulness of the proposed measure, we provide a descriptive comparison of TA between the groups of HEI. Despite the descriptive nature of the presented results, they can provide a basis for future research. For this purpose, we group the HEI in our data according to two antecedents discussed in section "Antecedents of TA": ownership and research orientation. Figure 1 provides an overview over the average TFs by type of HEI.

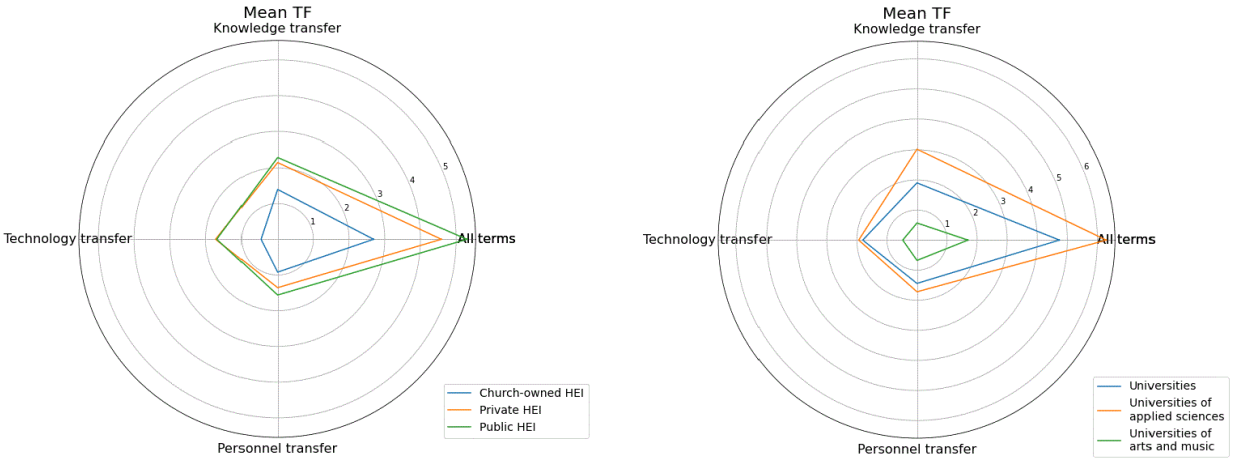

Figure 1. Average TF by research orientation and ownership

Comparing the TF, private and public HEI display very similar values, and indicate a higher presence of transfer of knowledge compared to other TA. If at all, public universities dedicate a slightly higher share of their communication to the overall topic of TA and its three dimensions on their websites. This supports the more recent view that public HEI are more successful regarding marketable TA (Bonaccorsi et al. 2022). Interestingly, church-owned HEI have on average a considerably lower tendency to communicate TA on their websites. They less often use terms related to TA (2.77 terms per 1000 compared to 4.62 or 5.36) and its categories. One striking difference is that church-owned HEI communicate relatively less frequently in the category of technology transfer (compared to the other categories), while transfer of personnel is the least displayed category for privately and state-owned HEI. This result is intuitive, as technology is closely tied to natural, physical or engineering sciences (Abreu and Grinevich 2013; O'Shea et al. 2005) as opposed to theology.

Concerning research orientation, our results suggest that "regular" universities of applied sciences (without the universities specialized in arts and music) exhibit considerably larger TF for all categories of TA. For example, regular universities use on average 4.74 per 1000 tokens related to TA and 1.90 per 1000 tokens related to transfer of knowledge, while universities of applied sciences use on average 6.29 per 1000 tokens for TA and 3.02 tokens per 1000 related to transfer of knowledge. This result is intuitive considering the different structures of those groups of HEI and the expectation that universities of applied sciences engage more actively in TA (Arvanitis et al. 2008; Jaeger and Kopper 2014). Moreover, universities of applied sciences seem to prefer transfer of knowledge over technology transfer according to their communication pattern, which supports the view that they engage especially in non-codified TA (Abreu and Grinevich 2013). Furthermore, universities for arts and music display substantially lower values for TF. Seeing that those HEI specialize in one field of education and research, this finding matches the view that specialist universities tend to have deficiencies in networks and opportunities necessary for TA (Bonaccorsi et al. 2022; Giuri et al. 2019). However, they seem to prefer non-codified TA (transfer of knowledge and transfer of personnel) over codified TA (technology transfer), which is again in line with the view of Abreu and Grinevich (2013).

Discussion

Contribution

This study combines inductive and deductive steps, as well as computational techniques and human input to analyze TA in German HEI. This comprehensive understanding of institutional TA contributes to broadening the research landscape in three interconnected ways.

First, we develop a new and open-source dictionary to measure HEI third mission activities. By means of inductive and deductive category building based on literature review, expertise and AI, we derive a comprehensive keyword list including a classification of activities towards the transfer of knowledge, technology, and personnel. The derivation and validation of the dictionary and its associated category system, as well as the described procedure, provide a valuable contribution for future researchers and acts as a foundation for the application of data science methods in AE and TA literature. Second, our proposed measure can help policymakers with the evaluation of HEI third mission activities. While these activities are institutionalized (Berghäuser 2017) and HEI are expected to pursue them complementary to their first two missions (Secundo et al. 2017), institutions are still seeking for appropriate evaluation methods (Etzkowitz 2016). To realize the increasing transfer orientation, HEI obtain funding from public sources, linked to specific objectives and performance agreements between the responsible ministries of the federal states and the respective HEI. This institutionalization raises the need for a development of indicators, allowing the measurement of TA on a HEI level. Thus, the proposed measure can assist governments in addressing this issue. Even though our study measures communicated activities, this approach can be seen as a first step that can serve as an evaluation basis for measuring TA. Third, our results underscore the significance of non-codified and multi-dimensional TA measurement empirically. Although prior research calls for the integration of more diverse TA (Perkmann et al. 2013), its presence has, until now, been perceived as anecdotal due to quantifiability challenges (Barbini et al. 2021). The methodology proposed in this study provides an important approach to address this dilemma in a diverse and widespread research field (D'Este and Patel 2007; Perkmann and Walsh 2007). Our results complement existing measures and help to minimize information deficits (Dörr et al. 2022). Although this is not the first study to apply web mining on HEI third mission activities, previous studies could not expose the relevance of TA in breadth and depth yet (Della Volpe and Esposito 2020). Our data shows that all HEI communicate about TA on a multi-dimensional level. Specifically, the results emphasize the importance of knowledge transfer, often missed in past TA research. Only universities for administrative sciences and universities of arts and music tend to communicate one other category more often, namely the transfer of personnel. The representative dataset, covering web content from 87.04% of all state-approved HEI in Germany, generates a large body of information that has not been used in this context before. It demonstrates the potential of web mining techniques compared to traditional quantitative methods such as questionnaires, which can be highly time consuming, cost inefficient and generate low response rates (e.g., Abreu and Grinevich 2013; Arvanitis et al. 2008; D'Este & Patel 2007). When controlling for TA antecedents, our findings are in line with previous studies (e.g., Abreu and Grinevich 2013; O'Shea et al. 2005). This reaffirms that web mining can serve as an efficacious instrument, especially when extracting non-numeric data. Interestingly, and in line with Bonaccorsi et al. (2022), our results also reveal that public universities are more active in communicating TA compared to private or church-affiliated ones. Universities with a practical focus, like those in applied sciences, are more engaged in TA than other types. Although these findings are intuitive, they indicate high validity of the proposed methodology.

Limitations and future research

As with every web mining-based paper, our study is not without limitations. First, our results depend strongly on the quality of the applied keywords. Languages are very complex and although they are easily understandable for humans, a computer cannot detect the meaning of a word as easily. It is important to note that although we include many relevant terms referring to TA, there is a risk of missing out specific names of TA programs that are unique in countries, states or even HEI. This limitation is potentially more severe for the English language. The challenge of constructing a list that contains all possible words related to TA that are at the same time easily distinguishable from other meanings remains complex. Finding a reference to TA that matches with our proposed keywords does not simultaneously mean that this activity is already undertaken or whether it is a commitment to engaging in such activities in the near future. Also,

some keywords might be used in other contexts, which are not related to TA directly. However, we apply several steps to reduce this bias. More precisely, the derivation and categorization of the keyword list is strongly rooted in the literature, furthermore we use AI and several manual control rounds to ensure the quality of our applied keywords.

Moreover, the results rely on self-reported website texts of the HEI. Activities that are not reported online cannot be detected and are therefore not represented in the results. On the other hand, there might be a bias due to over-reporting, if for example a cooperation project is advertised on more than one page of the website and using more than just one (or few) of the keywords in our dictionary. Hence, it is important to differentiate between communicated and conducted TA. Nevertheless, HEI are expected and somewhat forced to publicly communicate their TA due to an institutionalized legitimacy and accountability (Berman 2012; Fini et al. 2010; Murray 2010). Informing the public and stakeholders about their activities is contractually governed in the respective state law of each HEI (Berghäuser 2017), underlining communication of TA through websites as a suitable indicator for measuring TA.

The methodological focus of this study leads to a trade-off in terms of its theoretical impact or to follow the words of Sutton and Staw (1995: p.374): "Data are not theory". We acknowledge, that this is a limitation of our paper. However, we focus on a first application of web-mining techniques regarding TA in the German HEI context using a specially designed classification of TA. According to Gregor's (2006) taxonomy, the study can be classified as a "Theory for Analyzing". By integrating and measuring non-codified transfer activities, we try to find answers to the question of "what is" (Miles and Hubermann 1994), focusing on the transmitter perspective of knowledge: namely the HEI. Causal relationships or predictive generalizations contain potential for future research projects. Analyzing the receiver's perspective is a promising future research avenue, for example by investigating how and to what extent the TA spillover to regional economies, or societal stakeholders. Thus, our study provides an important starting point for future research in terms of classification and measurement of communicated TA, fulfilling a necessary condition for future research (Gregor 2006: p. 633).

Given the limited scope of our study, it is important to discuss the transferability of our proposed methodology and keyword list to other geographical contexts. Despite substantial public investment in education and research, European countries yield fewer economic benefits compared to nations like the USA. To address this, the European Union (EU) instituted the "Knowledge Triangle", integrating education, research, and innovation, thus leveraging universities as catalysts for economic growth (García et al. 2012). Initiatives like the Bologna Process (1999), the Lisbon Strategy (2000) and Horizon Europe (2020) highlight the critical role of R&D and education in the labor market (Maassen and Stensaker 2011), positioning universities as pivotal for EU economic competitiveness (Berghäuser 2017). Following these reforms, EU member states demonstrate shared political and educational commonalities, such as standardized teaching assessments (ECTS) or modular studies resulting in an approximation of study degrees. Since policymakers frequently adapt models from renowned institutions like Harvard or Stanford and effective political interventions such as the Bayh-Dole Act (Baldini et al. 2015; OECD 2003), the resulting legislations for HEI are aligned. Although some countries, like Ireland, Italy or Sweden, maintain distinct legislative frameworks regarding specific TA domains (i.e., IP rights) (Baldini et al. 2015), many other nations, including Denmark, Finland, France, Germany, Greece, Netherlands, Norway, Portugal or United Kingdom, have granted HEI rights to exploit research technologies, leading to similar TA infrastructures, like TTOs and entrepreneurship centers (Geuna and Muscio 2009). Hence, the global role of HEI in fostering economic, ecological, and social development is increasingly recognized, including lower and mid-income countries (Göransson et al. 2009). Thus, Germany's experience likely provides insights applicable to other EU states, given their shared policy and educational paths. Consequently, researchers can employ our methodology and keyword list for web mining in TA research across other regions. It remains essential, however, to tailor the keyword list to specific national attributes, ensuring comprehensive TA measurements. This adjustment enables future studies to evaluate TA's regional economic impact and unearth additional TA antecedents.

Considering the methodological focus of our study, we propose to see our results as first insights and a first step to a new indicator for TA. While the evaluation of the first two missions of HEI, namely teaching and research, is well-established—evidenced by studies examining imprinting effects via HEI alumni mobility (Klingler-Vidra & Chalmers 2023) or through patent and financial data (European Commission 2018)—there remains a gap in assessing third mission activities. Notably, while the transition of HEI alumni into

the industry is a prevalent form of personnel transfer (Gulbrandsen and Slipersoeter 2007), and publications serve as a medium for knowledge transfer, these metrics often overlook informal third mission activities with external partners, such as guest lectures, internships, or knowledge-sharing initiatives (García et al. 2012). Our proposed indicator does not offer a standalone evaluation of HEI TA. Instead, it provides an avenue for future research to integrate it with established measures, ensuring a comprehensive analysis of HEI.

Lastly, our results presented in this study can be used to train more elaborate machine learning models to track TA on websites from HEI, but also other actors (e.g., from the private sector) in the future. They can for example help in data cleaning by filtering only relevant texts (texts that contain at least one of the keywords) and thus reducing data size. Moreover, such filtering can improve the results of text classification as the algorithm is trained on more homogenous texts.

References

Abreu, M. & Grinevich, V. (2013). The nature of academic entrepreneurship in the UK: widening the focus on entrepreneurial activities. *Research Policy*, 42(2), 408-422.

Acs, Z. J., Stam, E., Audretsch, D. B. & O'Connor, A. (2017). The lineages of the entrepreneurial ecosystem approach. *Small Business Economics*, 49(1), 1-10.

Albahari, A., Barge-Gil, A., Pérez-Canto, S. & Landoni, P. (2022). The effect of science and technology parks on tenant firms: a literature review. *Journal of Technology Transfer*.

Alsos, G., Hägg, G., Lundqvist, M., Politis, D., Stockhaus, M., Williams-Middleton, K. & Djupdal, K. 2023. Graduates of venture creation programs – where do they apply their entrepreneurial competencies? *Small Business Economics*, 60, 133-155.

Aman, V. (2020). Transfer of knowledge through international scientific mobility: Introduction of a network-based bibliometric approach to study different knowledge types. *Quantitative Science Studies*, 1(2), 565-581.

Aizawa, A. (2003). An information-theoretic perspective of tf–idf measures. *Information Processing & Management,* 39(1), 45-65.

Arvanitis, S., Kubli, U. & Woerter, M. (2008). University-industry knowledge and technology transfer in Switzerland: What university scientists think about co-operation with private enterprises. *Research Policy*, 37(10), 1865-1883.

Azoulay, P., Ding, W. & Stuart, T. (2009). The impact of academic patenting on the rate, quality and direction of (public) research output. *The Journal of Industrial Economics*, 57(4), 637-676.

Backer, T. E., Pan, E. & Wolek, F. W. (1980). Network building conferences in support of industry-wide innovations. *Journal of Technology Transfer*, 4, 47-57.

Baena-García, M., Carmona-Cejudo, J. M., Castillo, G., & Morales-Bueno, R. (2011). TF-SIDF: Term frequency, sketched inverse document frequency. *2011 11th International Conference on Intelligent Systems Design and Applications,* 1044-1049, IEEE.

Baldini, N., Fini, R., & Grimaldi, R. (2015). The Transition Towards Entrepreneurial Universities: An Assessment of Academic Entrepreneurship in Italy. *SSRN Electronic Journal*.

Barbini, F.M., Corsino, M. & Giuri, P. (2021). How do universities shape founding teams? Social proximity and informal mechanisms of knowledge transfer in student entrepreneurship. *Journal of Technology Transfer*, 46, 1046–1082.

Beel, J., Langer, S., & Gipp, B. (2017). TF-IDuF: a novel term-weighting scheme for user modeling based on users' personal document collections.

Bektaş, Ç. & Tayauova, G. (2015). A model suggestion for improving the efficiency of higher education: University-industry cooperation. *Procedia - Social and Behavioral Sciences*, 191, 1235-1240.

Belenzon, S., & Schankerman, M. 2009. University Knowledge Transfer: Private Ownership, Incentives and Local Development Objectives. *Journal of Law and Economics*, 52(1), 111-144.

Berbegal-Mirabent, J., Ribeiro-Soriano, D. E., & García, J. L. S. (2015). Can a magic recipe foster university spin-off creation? *Journal of Business Research*, 68(11), 2272-2278.

Berghäuser, H. (2017). Die Dritte Mission in der Hochschulgovernance. Eine Analyse der Landeshochschulgesetze. Hochschulmanagement, 12(2-3), 35–43.

Berman, E. P. (2012). Explaining the move toward the market in US academic science: How institutional logics can change without institutional entrepreneurs. Theory and Society, 41(3), 261–299.

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.

Bischoff, K., Volkmann, C. K., & Audretsch, D. B. (2018). Stakeholder collaboration in entrepreneurship education: an analysis of the entrepreneurial ecosystems of European higher educational institutions. *Journal of Technology Transfer*, 43, 20-46.

Blank, T. H. (2021). When incubator resources are crucial: survival chances of student startups operating in an academic incubator. *Journal of Technology Transfer*, 46, 1845-1868.

Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-K text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, 50(4), 623-646.

Bonaccorsi, A., Blasi, B., Anna, C., & Sandra, N. (2022). Quality of research as source and signal: revisiting the valorization process beyond substitution vs complementarity. *The Journal of Technology Transfer*, 47, 407-434.

Boroditski, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition*, 118(1), 123-129.

Bun, K. K., & Ishizuka, M. (2001, June). Emerging topic tracking system. *In Proceedings Third International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems. WECWIS 2001,* pp. 2-11, IEEE.

Compagnucci, L., & Spigarelli, F. (2020). The Third Mission of the university: A systematic literature review on potentials and constraints. *Technological Forecasting and Social Change*, 161, 120284.

Crişan, E.L., Salanţă, I.I., Beleiu, I.N., Bordean, O.N. & Bunduchi, R. (2021). A systematic literature review on accelerators. *Journal of Technology Transfer*, 46, 62–89.

D'Este, P., & Patel, P. (2007). University–industry linkages in the UK: What are the factors underlying the variety of interactions with industry? *Research Policy*, 36(9), 1295–1313.

Davenport, T.H. & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press, Boston.

De La Torre, E. M., Agasisti, T., & Perez-Esparrells, C. (2017). The relevance of knowledge transfer for universities' efficiency scores: an empirical approximation on the Spanish public higher education system. *Research Evaluation*, 26(3), 211-229.

De Wit-de Vries, E., Dolfsma, W.A., van der Windt, H.J. & Gerkema, M.P. (2019). Knowledge transfer in university-industry research partnerships: a review. *Journal of Technology Transfer*, 44, 1236-1255.

Della Volpe, M., & Esposito, F. (2020). Discursive practices about third mission. A survey from Italian universities' official websites. *Quality in Higher Education*, 26(2), 224-239.

Dolmans, S.A.M., Walrave, B., Read, S., & van Stijn, N. (2022). Knowledge transfer to industry: How academic researchers learn to become boundary spanners during academic engagement. *The Journal of Technology Transfer*, 47(5), 1422–1450.

Dörr, J.O., Kinne, J., Lenz, D., Licht, G., & Winker, P. (2022). An integrated data framework for policy guidance during the coronavirus pandemic: Towards real-time decision support for economic policymakers. *PloS one*, 17(2), e0263898.

Dosi, G. (1982). Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11(3), 147-162.

Etzkowitz, H. (2016). The Entrepreneurial University: Vision and Metrics. *Industry and Higher Education*, 30(2), 83–97.

Etzkowitz, H. (1998). The norms of entrepreneurial science: cognitive effects of the new university–industry linkages. *Research Policy*, 27, 823–833.

Etzkowitz, H. (2003). Research groups as 'quasi-firms': the invention of the entrepreneurial university. *Research Policy*, 32(1), 109–121.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from national systems and 'Mode 2' to a triple helix of university-industry-government relations. *Research Policy*, 29(2), 109-123.

Etzkowitz, H., & Webster, A. (1998). Entrepreneurial science: the second academic revolution, in *Capitalizing knowledge: New intersections of Industry and Academia*, Etzkowitz, E., Webster, A., Healey, P. (eds.), Albany: SUNY Press.

European Commission (2018). *European Innovation Scoreboard 2018*. Publications Office.

Faccin, K., De Beer, C., Martins, B.V., Zanandrea, G, Kela, N. & Schutte, C. (2022). What really matters for TTOs efficiency? An analysis of TTOs in developed and developing economies. *Journal of Technology Transfer*, 47, 1135-1161.

Fagerberg, J., & Srholec, M. (2008). National innovation systems, capabilities and economic development. *Research Policy*, 37(9), 1417–1435.

Ferreira, J.M., & Carayannis, E. G. (2019). University-industry knowledge transfer - unpacking the "black box": an introduction. *Knowledge Management Research & Practice*, 17(4), 353-357.

Fini, R., Fu, K., Mathisen, M., Rasmussen, E., & Wright, M. (2017). Institutional determinants of university spin-off quantity and quality: a longitudinal, multilevel, cross-country study. *Small Business Economics*, 48(2), 361–391.

Fini, R., Lacetera, N., & Shane, S. (2010). Inside or outside the IP system? Business creation in academia. *Research Policy*, 39(8), 1060–1069.

García, A. C., Sierra, J. M. C., Soeiro, A., Hämäläinen, K., Boffo, S., Pausits, A., Murphy, M., Marhl, M., Vidal, J., Mora, J. G., & Padfield, C. (2012). *Green Paper. Fostering and Measuring ´Third Mission´ in Higher Education Institutions*. E3M Project.

Geuna, A., & Muscio, A. (2009). The Governance of University Knowledge Transfer: A Critical Review of the Literature. *Minerva*, 47(1), 93-114.

Giesel, K. D. (2007). *Leitbilder in den Sozialwissenschaften: Begriffe, Theorien und Forschungskonzepte* (1). VS Verlag für Sozialwissenschaften.

Giones, F., Kleine, K., & Tegtmeier, S. (2022). Students as scientists' co-pilots at the onset of technology transfer: a two-way learning process. *Journal of Technology Transfer*, 47, 1373-1394.

Giuri, P., Munari, F., Scandura, A., & Toschi, L. (2019). The strategic orientation of universities in knowledge transfer activities. *Technological Forecasting and Social Change*, 138, 261-278.

Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. In: *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*.

Göransson, B., Maharajh, R., & Schmoch, U. (2009). New activities of universities in transfer and extension: Multiple requirements and manifold solutions. *Science and Public Policy*, 36(2), 157–164.

Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3), 611-642.

Guiso, L., Sapienza, P., & Zingales, L. (2015). The value of corporate culture. *Journal of Financial Economics*, 117(1), 60-76.

Gulbrandsen, M., & Slipersoeter, S. (2007). The Third Mission and the Entrepreneurial University Model. In A. Bonaccorsi & C. Daraio (Eds.), *Universities and Strategic Knowledge Creation*. Edward Elgar Publishing.

Hayter, C.S. (2015). Public or private entrepreneurship? revisiting motivations and definitions of success among academic entrepreneurs. *Journal of Technology Transfer*, 40(6), 1003–1015.

Hayter, C.S., Rasmussen, E., & Rooksby, J.H. (2020). Beyond formal university technology transfer: innovative pathways for knowledge exchange. *Journal of Technology Transfer*, 45, 1-8.

Hess, S., & Siegwart, R.Y. (2013). R&D Venture: proposition of a technology transfer concept for breakthrough technologies with R&D cooperation: A case study in the energy sector. *Journal of Technology Transfer*, 38, 153-179.

Inkpen, A.C., & Tsang, E.W.K. (2005). Social Capital, Networks, and Knowledge Transfer. *The Academy of Management Review*, 30(1), 146–165.

Jaeger, A., & Kopper, J. (2014). Third mission potential in higher education: measuring the regional focus of different types of HEIs. *Review of Regional Research*, 34, 95-118.

Johnson, J.R. (1989). *Technology: Report of the Project 2061 Phase I technology panel*. AAAS Books.

Karnani, F. (2013). The university's unknown knowledge: tacit knowledge, technology transfer and university spin-offs findings from an empirical study based on the theory of knowledge. *Journal of Technology Transfer*, 38, 235-250.

Kindermann, B., Beutel, S., de Lomana, G.G., Strese, S., Bendig, D., & Brettel, M. (2021). Digital orientation: Conceptualization and operationalization of a new strategic orientation. *European Management Journal*, 39(5), 645-657.

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: a framework and a large-scale pilot study. *Scientometrics*, 125(3), 2011-2041.

Klingler-Vidra, R., & Chalmers, A. W. (2023). The Entrepreneurial University's Impact on Regional Socioeconomic Development: The "Alumni Policymaker" Mechanism. *Business and Politics*, 25(3), 251–292.

Lach, S., & Schankerman, M. (2004). Incentives and invention in universities. *The RAND Journal of Economics*, 45(1), 33-66.

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.

Lee, Y.S. (2000). The Sustainability of University-Industry Research Collaboration: An Empirical Assessment. *The Journal of Technology Transfer*, 25, 111-133.

Leitner, K.H., Bergner, S., & Rybnicek, R. (2021). The role of heads of departments in the commercialization of university research. *Journal of Business Economics*, 91, 353-378.
Link, A. N., Siegel, D. S., & Bozeman, B. (2007). An empirical analysis of the propensity of academics to engage in informal university technology transfer. *Industrial and Corporate Change*, 16(4), 641–655.
Lockett, A., Siegel, D., Wright, M., & Ensley, M. D. (2005). The creation of spin-off firms at public research institutions: Managerial and policy implications. *Research Policy*, 34(7), 981–993.
Maassen, P., & Stensaker, B. (2011). The knowledge triangle, European higher education policy logics and policy implications. *Higher Education*, 61(6), 757–769.
Marhl, M., & Pausits, A. (2013). Third mission indicators for new ranking methodologies. Lifelong Education: *The XXI Century*, 1(1), 89–101.
Mariani, G., Carlesi, A., & Scarfò, A. (2018). Academic spinoffs as a value driver for intellectual capital: The case of the University of Pisa. *Journal of Intellectual Capital*, 19(1), 202-226.
McKenny, A. F., Short, J. C., Ketchen Jr, D. J., Payne, G. T., & Moss, T. W. (2018). Strategic entrepreneurial orientation: Configurations, performance, and the effects of industry and time. *Strategic Entrepreneurship Journal*, 12(4), 504-521.
Menzies, T. V. (2000). An Exploratory Study of University Entrepreneurship Centres in Canada: A First Step in Model Buildings. *Journal of Small Business & Entrepreneurship*, 15(3), 15-38.
Meoli, M., Paleari, S., & Vismara, S. (2019). The governance of universities and the establishment of academic spin-offs. *Small Business Economics*, 52, 485-504.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE.
Molas-Gallart, J., Salter, A., Patel, P., Scott, A. & Duran, X. (2002). *Measuring Third Stream activities. Final Report to the Russell Group of Universities*. SPRU Science and Technology Policy Research Unity, University of Sussex, Brighton, UK.
Montesinos, P., Carot, J. M., Martinez, J., & Mora, F. (2008). Third Mission Ranking for World Class Universities: Beyond Teaching and Research. *Higher Education in Europe*, 33(2–3), 259–271.
Morandi, V. (2013). The management of industry-university joint research projects: how do partners coordinate and control R&D activities? *Journal of Technology Transfer*, 38, 69-92.
Moss, T. W., Renko, M., Block, E., & Meyskens, M. (2018). Funding the story of hybrid ventures: Crowdfunder lending preferences and linguistic hybridity. *Journal of Business Venturing*, 33(5), 643-659.
Mowery, D. C., Nelson, R. R., Sampat, B. et al. (2004). *Ivory Tower and Industrial Innovation: University-Industry Technology Before and After the Bayh-Dole Act in the United States*. Stanford University Press, California.
Murray, F. (2010). The Oncomouse That Roared: Hybrid Exchange Strategies as a Source of Distinction at the Boundary of Overlapping Institutions. *American Journal of Sociology*, 116(2), 341–388.
Muscio, A., Shibayama, S., & Ramaciotti, L. (2022). Universities and start-up creation by Ph.D. graduates: the role of scientific and social capital of academic laboratories. *The Journal of Technology Transfer*, 47, 1-29.
O'Carroll, C., Harmon, C., Farrell, L. (2006). *The Economic and Social Impact of Higher Education*. Irish Universities Association, Dublin.
O'Shea, R. P., Chugh, H., Allen, T. J. (2008): Determinants and consequences of university spinoff activity: a conceptual framework. *The Journal of Technology Transfer*, 33(6), 653–666.
O'Shea, R. P., Allen, T. J., Chevalier, A. et al. (2005): Entrepreneurial orientation, technology transfer and spinoff performance of U.S. universities. *Research Policy*, 34(7), 994-1009.
Obschonka, M., Lee, N., Rodríguez-Pose, A. et al. (2020). Big data methods, social media, and the psychology of entrepreneurial regions: capturing cross-county personality traits and their impact on entrepreneurship in the USA. *Small Business Economics*, 55(3), 567-588.
OECD (2003). *Turning science into business: Patenting and licensing at public research organisations. Organisation for Economic Co-Operation and Development*. OECD Publishing.
Palmer, D. T. (2010). Knowledge Exchange, A Strategy for Open Access Success. *Chinese Journal of Library and Information Science*, 4(1), 81-94.
Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., & Sobrero, M. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy*, 42(2), 423–442.
Perkmann, M., & Walsh, K. (2007). University–industry relationships and open innovation: Towards a research agenda. *International Journal of Management Reviews*, 9(4), 259–280

Phan, P., & Siegel, D. S. (2006). The effectiveness of university technology transfer. *Foundations and Trends in Entrepreneurship*, 2(2), 77–144.
Prüfer, J., & Prüfer, P. (2020). Data science for entrepreneurship research: studying demand dynamics for entrepreneurial skills in the Netherlands. *Small Business Economics*, 55(3), 651-672.
Rasmussen E., Mosey S., Wright M. (2015). The transformation of network ties to develop entrepreneurial competencies for university spin-offs. *Entrepreneurship & Regional Development*, 27(7/8), 430–457.
Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of documentation*, 60(5), 503-520.
Rolfo, S., & Finardi, U. (2014). University third mission in Italy: organization, faculty attitude and academic specialization. *Journal of Technology Transfer*, 39(3), 472-486.
Sapir, E. (1944). Grading, a study in semantics. *Philosophy of science*, 11(2), 93-116.
Schaeffer, V., Öcalan-Özel, S., & Pénin, J. (2020). The complementarities between formal and informal channels of university-industry knowledge transfer: a longitudinal approach. *Journal of Technology Transfer*, 45(1), 31-55.
Schultz, C., Gretsch, O., & Kock, A. (2021). The influence of shared R&D-project innovativeness perceptions on university-industry collaboration performance. *Journal of Technology Transfer*, 46(3), 1144-1172.
Secundo, G., Elena Perez, S., Martinaitis, Ž., & Leitner, K. H. (2017). An Intellectual Capital framework to measure universities' third mission activities. *Technological Forecasting and Social Change*, 123, 229–239.
Short, J. C., Broberg, J. C., Cogliser, C. C., & Brigham, K. H. (2010). Construct validation using computer-aided text analysis (CATA) an illustration using entrepreneurial orientation. *Organizational Research Methods*, 13(2), 320-347.
Siegel, D. S. (2006). *Technology Entrepreneurship: Institutions and Agents Involved in University Technology transfer*. Edgar Elgar, London.
Siegel, D. S., Waldman, D., & Link, A. (2003). Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study. *Research Policy*, 32(1), 27-48.
Siegel, D. S., Veugelers, R., Wright, M. (2007). Technology transfer offices and commercialization of university intellectual property: performance and policy implications. *Oxford Review of Economic Policy,* 23(4), 640-660.
Sutton, R. I., & Staw, B. M. (1995). What Theory is Not. *Administrative Science Quarterly*, 40(3), 371-384.
Svensson, P., Klofsten, M., & Etzkowitz, H. (2012). An entrepreneurial university strategy for renewing a declining industrial city: the Norrköping way. *European Planning Studies*, 20(4), 505-525.
Ulrich, P., & Fluri, E. (1995). *Management: Eine konzentrierte Einführung* (7). Haupt.
Urdari, C., Farcas, T., & Tiron Tudor, A. (2017). Assessing the legitimacy of HEIs' contributions to society: the perspective of international rankings. Sustainability Accounting. *Management and Policy Journal*, 8(4), 471-492.
Vaupel, M., Bendig, D., Fischer-Kreer, D., & Brettel, M. (2022). The Role of Share Repurchases for Firms' Social and Environmental Sustainability. *Journal of Business Ethics*, 1-28.
Verspagen, B. (2006). Innovation and Economic Growth. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *The Oxford Handbook of Economic Growth* (pp. 487-513). Oxford University Press.
Vinig, G. T. & van Rijsbergen, P. (2009). Determinants of University Technology Transfer - Comparative Study of US, Europe and Australian Universities. *SSRN*.
Vorley, T., & Nelles, J. (2008). (Re)Conceptualising the academy: institutional development of and beyond the third mission. *Higher Education Management and Policy*, 20(3), 1-17.
Ward, V., House, A., & Hamer, S. (2009). Developing a framework for transferring knowledge into action: A thematic analysis of the literature. *Journal of Health Services Research & Policy*, 14(3), 156-164.
Whorf, B. L. (1956). Language, thought, and reality: selected writings of….(Edited by John B. Carroll.)
Wright, M. (2014). Academic entrepreneurship, technology transfer and society: Where next? *The Journal of Technology Transfer*, 39(3), 322–334.
Wright, M., Birley, S., & Mosey, S. (2004). Entrepreneurship and University Technology Transfer. *The Journal of Technology Transfer*, 29(3/4), 235–246.
Zhao, Z., Broström, A., & Cai, J. (2020). Promoting academic engagement: university context and individual characteristics. *Journal of Technology Transfer*, 45(1), 304–337.