Bridging Robotics/AI and Real-World Labs: A Quantitative Approach Based on Mining German Newspaper Articles

Martha Loewe[a], Ingrid Ott[a,b,c,*]

[a] KIT, Institute of Economics, Chair of Economic Policy, Kaiserstr. 12, Karlsruhe, 76131, Germany
[b] Kiel Institute for the World Economy (IfW), Kiellinie 66, Kiel, 24105, Germany
[c] CESifo, Poschingerstr. 5 , Munich, 81675, Germany

ARTICLE INFO

Keywords:
real world lab
regulatory sandbox
smart regulation
service robotics
media analysis
structural topic modeling
Declaration of interest: none.

ABSTRACT

Both robotics/AI (RAI) and real world labs (RWL) are current topics in public innovation promotion policies, but are mostly treated in isolation. While RAI has a focus on a specific technology to serve society, RWL address the institutional context including experimental learning of governments and societal perspectives. We are particularly interested in the interface between RAI and RWL and the way media is reporting on these two domains. This reflects key aspects of the social debate in relation to RAI and RWL. We base our analysis on the understanding that technology development and diffusion ultimately depend on institutional arrangements that are developed alongside or in lieu of market arrangements and also reflect societal needs. This paper uses quantitative text analysis to examine 3,800 German broadsheet newspaper articles in the period 2016-2023. We use Structural Topic Modeling (STM) with publication date and sub-corpus source as covariates to trace topic dynamics and topical prevalence contrast. We show that the dominant topic has changed over time from RAI ("Machine Learning and AI Development Methods") to RWL ("Real-World Labs for the Energy Transition"). We identify bridge topics and argue that these are diverse and include philosophical and legal considerations, public funding and specific application areas for robots, e.g. in schools. As indicators to identify the interface between the two domains (RAI, RWL), we propose a combination of topical prevalence contrast and eigenvector centrality and the use of psycholinguistic attributes to evaluate the topics. These elements could be broadly used to exploit possible complementarities for government experimental learning and when designing "smart regulation" which targets several fields simultaneously.

Acknowledgements

We thank Barbara Bruno, Franziska Krebs, Moritz Müller, Utku Norman and Nora Weinberger for their helpful feedback on an earlier draft of this paper, Jonathan Völkle for his excellent research assistance, and Tamin Asfour for his ongoing support.

CRediT authorship contribution statement

Martha Loewe: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing Original Draft, Review & Editing, Visalization. *Ingrid Ott*: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Writing Original Draft, Review & Editing.

1. Introduction

Both technology development and diffusion ultimately depend on institutional arrangements that are developed alongside or in lieu of market arrangements and also reflect societal needs.

Questions arise concerning the optimal institutional design which enables not only the development of technology, but also its diffusion. In line with the systems of innovation literature, the aim is to include the perspective of as many actors as possible in the analysis. In times of multiple crises and rapid change in an evermore complex world, new technologies, processes and policies are increasingly tested in designated spaces and scales under real conditions. These test rooms mostly refer to research and experimentation at the interface of science and society where primarily

martha.loewe@partner.kit.edu (M. Loewe); ingrid.ott@kit.edu (I. Ott)
wipo.kit.edu (M. Loewe); wipo.econ.kit.edu (I. Ott)
ORCID(s): 0000-0000-4515-2871 (I. Ott)

Highlights

Bridging Robotics/AI and Real-World Labs: A Quantitative Approach Based on Mining German Newspaper Articles

author1,author2

- We analyze 3800 German newspaper articles on robotics/AI and real world labs with STM.
- We conceptualize overlaps of usually separately analyzed fields as bridge topics.
- This contributes to the evolution of evidence-based instruments of smart regulation.
- Smart regulation develops policy choices for innovation systems minding the public.
- We offer insights to the stance of the general public based on measurable evidence.

Bridging Robotics/AI and Real-World Labs: A Quantitative Approach Based on Mining German Newspaper Articles

author1[a], author2[a,b,c,*]

[a] affiliation1,
[b] affiliation2,
[c] affiliation3,

ARTICLE INFO

Keywords:
real world lab
regulatory sandbox
smart regulation
service robotics
media analysis
structural topic modeling
Declaration of interest: none.

ABSTRACT

Both robotics/AI (RAI) and real world labs (RWL) are current topics in public innovation promotion policies, but are mostly treated in isolation. While RAI has a focus on a specific technology to serve society, RWL address the institutional context including experimental learning of governments and societal perspectives. We are particularly interested in the interface between RAI and RWL and the way media is reporting on these two domains. This reflects key aspects of the social debate in relation to RAI and RWL We base our analysis on the understanding that technology development and diffusion ultimately depend on institutional arrangements that are developed alongside or in lieu of market arrangements and also reflect societal needs. This paper uses quantitative text analysis to examine 3,800 German broadsheet newspaper articles in the period 2016-2023. We use Structural Topic Modeling (STM) with publication date and sub-corpus source as covariates to trace topic dynamics and topical prevalence contrast. We show that the dominant topic has changed over time from RAI ("Machine Learning and AI Development Methods") to RWL ("Real-World Labs for the Energy Transition"). We identify bridge topics and argue that these are diverse and include philosophical and legal considerations, public funding and specific application areas for robots, e.g. in schools. As indicators to identify the interface between the two domains (RAI, RWL), we propose a combination of topical prevalence contrast and eigenvector centrality and the use of psycholinguistic attributes to evaluate the topics. These elements could be broadly used to exploit possible complementarities for government experimental learning and when designing "smart regulation" which targets several fields simultaneously.

1. Introduction

Both technology development and diffusion ultimately depend on institutional arrangements that are developed alongside or in lieu of market arrangements and also reflect societal needs.

Questions arise concerning the optimal institutional design which enables not only the development of technology, but also its diffusion. In line with the systems of innovation literature, the aim is to include the perspective of as many actors as possible in the analysis. In times of multiple crises and rapid change in an evermore complex world, new technologies, processes and policies are increasingly tested in designated spaces and scales under real conditions. These test rooms mostly refer to research and experimentation at the interface of science and society where primarily solutions are sought for societal challenges and transformation processes. There are various terms for this emerging research format, including "real-world lab" (RWL) and "regulatory sandbox".

According to the German Federal Government, real-world laboratories as regulatory sandboxes are regarded as an effective means for developing innovation-friendly framework conditions and as spaces to test the impact of various forms of small-scale regulations.

The "learning" element in this format refers to many stakeholders, explicitly including the government, and the special role of society as a driver or inhibitor in the innovation process is emphasized. As with any new policy instrument, there is a need to review its impact. Ideally, it should fit seamlessly into existing contexts and mutually compatible funding instruments should be developed. There are (at least) two main challenges in this context. First, RWLs are heterogeneous and therefore context-specific. Given their focus on exploratory approaches and often small-scale and context-sensitive settings, regulatory sandboxes naturally fail to provide the indisputable evidence needed to

author1 (author1); email author2 (author2)
URLauthor1 (author1); URLauthor2 (author2)
ORCID(s):

motivate and defend government regulation at large. Second, quantifying the innovation induced by RWLs is a major challenge. There is a lack of data, and suitable indicators still need to be developed. While there are a large number of well-established input/output indicators that reflect innovation activities of companies (e.g. patents, publications, expenditure on research and development (R&D) personnel), it is much more difficult to find suitable indicators to measure the perspective of society.

Today, RWLs can be found in a wide variety of contexts. Most of them address topics related to sustainability focusing on the energy transition or mobility. However, there are a few RWLs that have a technical focus, including the "Real-World Lab Robotics Artificial Intelligence" at the Karlsruhe Institute of Technology (RAI; more below), that investigates the determinants of the innovation adoption and diffusion processes of humanoid robots into everyday life and public spaces on the basis of specific fields of application. The potential for automation in the service sector is still high and it is desirable to exploit this potential, especially in light of a shortage of skilled workers.One major challenge is the heterogeneity of the fields of robot application. In addition to the technological challenges, the area of data protection, for example, poses a major challenge. The question is how to organize (regulate) innovation processes intelligently so that they are legally compliant on the one hand and thus provide planning security for innovating companies and on the other hand allow sufficient freedom to experiment with a wide variety of formats and thus enable innovation in the first place. For innovation diffusion, the robots need to be adapted given specific constraints and individual requirements. It is not yet clear exactly what these are, and a RWL is a natural format to test them.

This paper combines the two perspectives of technological and institutional development while paying special attention to society. It analyzes how the two domains of RWL and RAI are linked in the media discourse that both informs and is shaped by society. The discussions thus embody the attitudes of the general public. We also focus on the representation of the government in this discourse. The joint consideration of RAI and RWL supports in searching key overlaps and solutions to two major challenges that otherwise are mostly analyzed in isolation. For a targeted design of policy instruments, it is also helpful to understand the attitudes of the general public towards certain topics. Especially in times of tight budgets, it is crucial to leverage synergies between different funding formats.

Methodologically, we use a method of quantitative content analysis, namely an unsupervised machine learning approach, called structural topic modeling (STM), that allows us to take into account exogenous covariates like the text source and the publication date of the articles. Our data consists of about 3,800 articles from German broadsheet newspapers published in the period from January 2016 to June 2023, reflecting a broad political spectrum. Newspapers are among the less biased sources for the study of evolving dynamics when monitoring trends and changes over time. We show that there is a variety of themes within our text corpus, ranging from philosophical considerations and science fiction over work place implications, specific application areas, machine learning methods, start-ups, funding, the energy transition and mobility. Structural topic modeling structures the collection of articles into topics automatically. However, the labeling of the topics and the interpretation of the results are based on extensive manual work and require critical reflection. We partition the entire text corpus into 32 topics and obtain the following results. First, the 32 topics can be assigned to the areas RWL or RAI to varying degrees, including bridge topics that are related to both dimensions RWL and RAI. We consider these bridge topics to be particularly interesting against the background of joint perspectives on funding policy or regulation that could be more broadly based. Second, we create a network that represents the structure of the relationships between the topics and highlights which of them are often discussed together within a newspaper article. We apply simple network indicators and uncover that "Philosophical Considerations on the Digital Transformation" followed by "Political Support for AI and Robotics in Germany" gain the most prestige within the network. Third, we apply a sophisticated sentiment analysis that evaluates the topics based on the four psycholinguistic attributes arousal, valence, abstractness and imageability. We find that, disregarding the timeline, dominant (i.e. large and prestigious) topics have a positive connotation and are less abstract. Fourth, the proposed approach facilitates zooming in on different broader themes. In the field of mobility, for example, we can show which facets of the media discourse have a positive connotation and which are presented more negatively[1]. A differentiated analysis is also possible in the areas of "future work" and robot assistance systems (in the private household, at school).

This paper contributes to a better understanding of the attitudes of the general public based on quantitative indicators. It especially highlights the bridge topics as a potential element of smart regulation. Such an element of smart regulation can add further important perspectives for the development of flexible, but well-founded and data-based regulation and policy instruments for innovation promotion beyond the classic innovation indicators. Our analysis contributes to implementing smart regulation based on quantitative evidence of the attitudes of the general public as

[1]Sustainability related reporting is positively connotated while reporting on autonomous driving is negatively connotated, as texts on autonomous mobility tend to focus on accidents.

reflected by the media reporting. Ideally, smart regulation increasingly integrates the attitudes of the general public while developing public policy choices for innovation systems.

The remainder of this paper is organized as follows. We begin with an introduction of the building blocks of this paper, namely robotics AI and real-world labs and reflect on some related literature in section 2. In section 3, we briefly introduce the applied method, namely Structural Topic Modeling, and explain in detail our data sources, the preprocessing steps and the model specification. In section 4, we present the main results regarding labeling, dynamics, topic prevalence contrast, correlation and sentiment analysis, and discuss these results. Section 5 critically reflects on the methodology used and mentions some policy implications while section 6 concludes.

2. Building Blocks and Related Literature

AI is broadly accepted as being today's most important General Purpose Technology (GPT). This comes with some benefits (increase in overall productivity and thus increased global welfare) and some challenges (requirements to adjust well-established procedures and boundary conditions which is both time consuming and costly). According to Bresnahan and Trajtenberg (1995), who coined the term of GPTs, these are characterized by pervasiveness (they are used as inputs by many downstream sectors), inherent potential for technical improvements and innovational complementarities (meaning that the productivity of R&D in downstream sectors increases as a consequence of innovation in the GPT).

Bresnahan and Trajtenberg (1995) also already point to implications regarding reorganizations of well-established practices and work arrangements not only in the application sectors but even beyond. In doing so they implicitly already address discussions that today are framed in the context of social innovation though not labeling them accordingly. Another point already addressed is the need to co-design of institutional and organizational arrangements to fully exploit the welfare potential of GPTs thereby also affecting the present and future pace of innovation. Bekar, Carlaw and Lipsey (2018) apply and evolutionary approach and argue that GPTs transform the structure of the economy and today's knwoledge society - where the knowledge in part is also gained via media. However, this literature does not address issues around regulatory learning but instead, for given boundary conditions, sees the government as a benevolent social planner that pursues the goal to internalize prevailing horizontal and vertical externalities to maximize overall welfare. One might conclude that the complementary perspective between economy, government and society is already implicitly included in their reasonings though not yet clearly spelled out.

Such a joint consideration of innovation and policy conditions, including the economic and the societal perspective, is is a matter of course nowadays (compare e.g. Grubb, Drummond, Poncia, McDowall, Popp, Samadi, Penasco, Gillingham, Smulders, Glachant and et al. (2021)). In particular, it has also been recognized that the co-evolution between technology development and institutional design requires the state, just like innovating companies, to test, evaluate and continuously develop its instruments. This is where RWLs come into play as modern tools.

In 2019 the German Federal Ministry for Economic Affairs and Energy defined real-world labs as "test spaces for innovation and regulation" für Wirtschaft und Energie (2019). Since 2018, the federal government has been prominently promoting the format of real-world laboratories as an explicit instrument for innovation, while there was already varying degrees of support at the level of individual federal states before that. Regulatory sandboxes, understood as research settings for developing, testing and evaluating solutions to societal problems, play an important role in the development of technologies that meet societal needs.

In 2020, the Council of the EU takes a position and concludes on regulatory sandboxes and experimentation clauses as instruments for an innovation-friendly, future-proof and resilient regulatory framework to address disruptive challenges in the digital age. The federal government in Germany is currently (as of March 2024) working on the adoption of a so-called real laboratory law (Reallabor-Gesetz). The key points here are defining overarching standards for RWLs, legal foundations for new RWLs in important areas of innovation, experimental clause check and a one-stop-shop for RWLs as a central point of contact for practice and knowledge transfer.

Today, within the scientific literature, RWL are seen as modern representations of innovation systems (compare e.g. Ott (forthcoming)). Regarding the literature on innovation systems, the paper at hand is most closely related to Technical Innovation Systems (TIS).[2] The TIS concept is concerned with the emergence of novel technologies and can be traced back to the seminal paper of Carlsson and Stankiewicz (1991). Research on TIS already includes the institutional and organizational changes that in addition to technology-push and demand-pull perspectives are seen as essential drivers behind the generation, diffusion, and utilization of technological innovation (compare Ott (forthcoming) for a

[2] Other prominent perspectives include a focus on specific sectors or on a special spatial range (regional, nation, global).

recent overview). RWLs are not limited in their applications and can represent technological, spatial and/ or topical foci thereby challenging the identification of overarching similarities of this new experimentation format (compare Schäpke, Stelzer, Caniglia, Bermann, Wanner, Singer-Bodowski, Loorbach, Ollson, Baedecker and Lang (2018)).

Regulatory sandboxes allow researchers or innovating firms to interact with diverse stakeholder groups to co-design and test socio-technical solutions. This is done by creating a less-regulated environment for a certain time period. Ideally, this contributes to gaining a deep understanding of the psychological and social processes affected by technological innovation and of user preferences, to explore desired and undesired effects of technology, but also to inspire, design and especially also to test the efficacy of policy tools. Despite these ambitious large-scale requirements, regulatory sandboxes are a highly context-sensitive and still a novel approach to studying socio-technical co-adaptation. Put differently, due to their context-specificity it is not easy to delineate conclusions that are valid in a larger context.

Currently, most RWLs have a strong environmental focus and only a few address the diffusion perspective in the context of technologies. Due to increasing computing capacity and miniaturization, modern robots are now flexible, easy to operate and becoming able to navigate autonomously, even in unstructured environments. Rapidly decreasing costs together with undeniable shortages of skilled workers in many fields (e.g. in the areas of elderly care, in education or in hospitals in general), are important drivers of the diffusion of so-called 'service robots'. These are robots performing useful tasks for humans or equipment excluding industrial applications.[3] The diffusion of service robotics today exceeds those of industrial robots and the International Federation of Robotics sees dominating future market potentials in that diverse field (compare IFR (2023)). However, SR are not a new phenomenon and their evolution can be traced technologically in patent data (compare Savin, Ott and Konop (2022)). At the same time, the technology is very heterogeneous in terms of the complexity, possible applications and prices which makes one-size-fits-all considerations impossible. In order to better understand their application potentials, it is therefore necessary to take into account the respective technological possibilities, the needs of customers and the regulatory environment. While the potential and the requirements for industrial robots are generally clearly specified, the situation is different with service robots, where the full application potential as well as diffusion potentials and obstacles are often not clear ex ante.[4] This is where the RWL RAI comes in.

The "Real-World Lab Robotics Artificial Intelligence" at the Karlsruhe Institute of Technology (KIT) operates at the interface between real-world labs on the one hand and technological development and innovation in the fields of robotics and AI on the other hand (Nierling, Weinberger, Vetter, Maia, Asfour, Krebs, Peller-Konrad, Reister, Younes, Bärmann and Loewe (2023)). The real-world lab deploys humanoid robots to various settings in the public square, including day care centers (Krebs, Peller-Konrad, Younes, Reister, Bärmann, Vetter, Weinberger, Loewe, Ott, Nierling and Asfour (2023); Rudenko, Norman, Maure, Rudenko, Weinberger, Krebs, Peller-Konrad, Asfour and Bruno (2024)), schools and museums, and gives members of the general public the opportunity to interact with them in the context of diverse experiments. Through these experiments, accompanying research, and citizens' dialogues, the researchers gain new insights into the preferences, expectations, desires and fears of potential users and take these observations into account while developing the design of the next generation of AI robots. Interacting with individuals is at the heart of this approach to exploring the attitudes of the general public. It inspired us to take an integrative look at two major funding lines of the German government, which were set up independently of each other and are usually analyzed independently of each other. The present study complements the work of "Robotics AI" by taking a bird's eye view. The aim is to develop perspectives on how experimental learning can be used to combine previously unconnected perspectives. also from the point of view of experimental governmental learning. In this way we make a contribution to a possible design of smart regulation (understood as the examination of forward-looking regulatory approaches and forms that transcend disciplinary and sectoral boundaries and pose questions for the future) and especially to quantifying some elements of it. The latter argument addresses a major shortcoming in the evaluation of RWLs, on the one hand, and in the scaling of insights, on the other.

In this paper, we use the term "real-world lab" in the widest possible sense, including *all* alternative definitions, as the newspaper articles that we analyze are not concerned with precise definitions, but with the concept as such. Other terms for similar concepts include "urban lab", "innovation lab", "future lab", or "transformation lab". Our

[3]For a precise definition of service robots and further sub-classifications, compare ISO 8371:2012, 2.11 (private use; synonyms are personal or domestic use) and 2.13. (professional use).

[4]Due to the multitude of forms and structures as well as application areas of service robots, it is sometimes not easy to delimit SR from industrial robots. E.g., in logistics, robots are used in non-manufacturing environments, such as logistic centers, hospitals or warehouses but also to transport parts within factories.

analyses have a technical RWLs in mind and allow us to simultaneously address the joint development of technology and institutions. As aforementioned we contextualize this for robotics and AI.

For our analyses, we take advantage of the fact that the strong link between innovation and economic, political, and socio-cultural factors is expressed in public media discourse. Despite the immense growth of social media and the associated importance it has as a source of information for many people (compare DellaVigna and Ferrar (2015)), newspapers are among the less biased sources for the study of evolving dynamics when monitoring changes and trends over time. It is also well recognized that individuals update their expectations when new information becomes available. The media can thus be collectively viewed as a suitable representation of the contents of discourses, notwithstanding that it might be biased (compare e.g. Gentzkow and Shapiro (2010), Lehotský, Černoch, Osička and Ocelík (2019) or more recently Cage, Hengel, Hervé and Urvoy (2024))). The media encompass a body of constantly evolving ideas and concepts that are generated, replicated, and adapted into practices that shape our understanding of reality. They serve as a proxy for discourse content, collectively acting as an arena for claims-making competition.

The evaluation of topics discussed in the media is increasingly (partially) automated. Due to the rising power of modern computers, the rising availability of large amounts of data and the further development of unsupervised machine learning methods, text as data has become a prominent source for analysis (e.g. Gentzkow and Shapiro (2010), Kelly, Papanikolaou, Seru and Taddy (2021), for a critical reflection see Grimmer and Stewart (2013)). STM is an unsupervised machine learning method to extract information from (large) textual data. Today, the use of STM by social scientist is exploding. The use of STM allows to add meta data, such as time or other covariates, e.g. the different text sources, to the text data and exploit that information for tracing trends and relationships between topics. STM as method has already been applied to in a variety of contexts, just to mention some compare e.g. regarding the relationship between science and technology in nanotechnology (compare Kang, Yang, Lee, Seo and Lee (2023)), related to drinking water quality (compare Sohns (2023)) or regarding media analyses (compare Lehotský et al. (2019) or Loewe, Quittkat, Knodt and Ott (2024)).

This study takes the bird's-eye view on society and analyzes the main themes in public reporting and discussion about robotics, AI and real-word labs. Policy conditions include the cultural and societal environment, that latter being reflected in the public discourse in newspapers, social networks, television, position papers and public speeches. However, a solid and state-of-the-art quantification of these phenomena is still missing. A key novelty of the paper is explicit bridging the perspectives of two important fields that usually are discussed, analyzed and also supported as isolated topics. We see that this may have important implications for governmental learning when thinking about the design of policy instruments and governmental funding schemes.

3. Data and Methods

3.1. Structural Topic Models (STMs) in a Nutshell

This study is based on 3,801 German newspaper articles that were published from January 1, 2016 to June 30, 2023. We evaluated the content of the collection of newspaper articles using a structural topic model (STM; compare Roberts, Stewart, Tingley, Lucas, Leder-Luis, Gadarian, Albertson and Rand (2014); Roberts, Stewart and Airoldi (2016a); Roberts, Stewart and Tingley (2016b)), a natural language processing method for automated content analysis. STMs are increasingly used to analyze large texts; studies closest to this paper are Dehler-Holland, Okoh and Keles (2022), Agrawal, Wankhede, Kumar, Luthra, Majumdar and Kazancoglu (2022), Zhang, Cao, Ji, Gu and Wang (2022) and Loewe et al. (2024).

Topic models in general uncover latent topics in a corpus (collection of documents) and structural topic models in particular facilitate assessing the influence of metadata on the topics. The STM is an extension of the Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan (2003)) and the Correlated Topic Model (CTM) (Blei and Lafferty (2007)). The advantage of the STM over other topic models is the option to add metadata information about the documents in the estimation of the model. We included two covariates: the publication date and an indicator variable that shows whether the article belongs to the Robotics/AI corpus or to the Real-World Lab corpus. The resulting model offers insights into the main themes in the German discourse about robotics/AI and real-world labs from 2016 to mid 2023 and sheds light on the linkages between the two domains.

The STM is a probabilistic model based on the bag-of-words approach, where documents are mapped to a distribution of their words, whereas the syntactical structure and the order of the words are disregarded. The algorithm partitions the distribution of words in the corpus into k topics. Topic models are mixed-membership models: they assume that each document is a combination of several topics with varying proportions, documents are not attributed

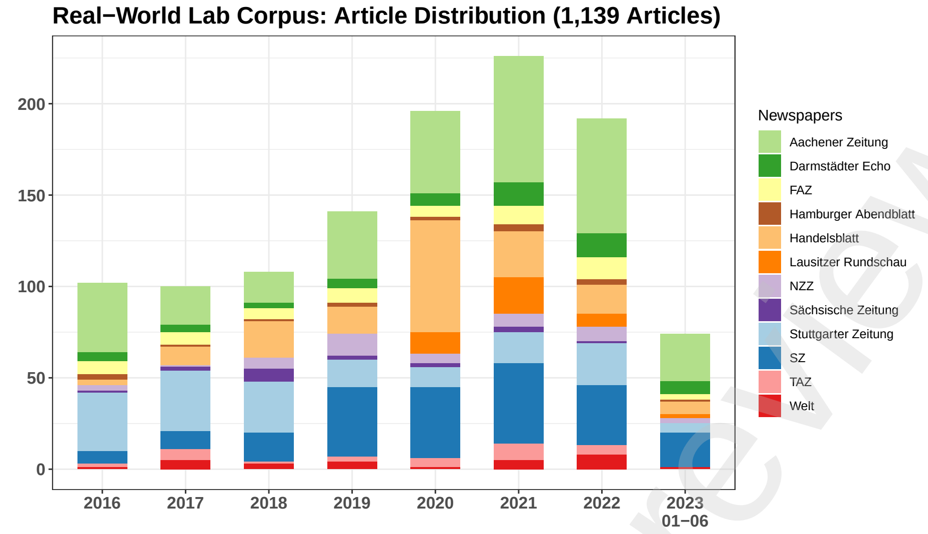


Figure 1: The distribution of 1,139 newspaper articles published in twelve German newspapers from January 1, 2016 to June 30, 2023 that feature the keyword "Reallabor" or one of the related terms. The related terms are given in Table 1.

to just a single topic. Topic models are unsupervised: given the number of topics k, they create a partition of the distribution of the words in the corpus into k parts based on variational inference, without any prior thematic input of the modeler. This method is particularly suited for exploratory research with limited a priori assumptions.

3.2. Data

Our collection of newspapers (corpus) consists of two sub-collections, related to real-world labs and robotics/AI respectively. The initial search for newspaper articles for the Real-World Lab corpus focused on articles in the largest German broadsheet newspapers that are read nation-wide, including *Süddeutsche Zeitung (SZ)*, *Frankfurter Allgemeine Zeitung (FAZ)*, *Die Tageszeitung (TAZ)*, *Die Welt*, *Handelsblatt* and *Neue Zürcher Zeitung (NZZ)*. This set of newspapers covers the political spectrum from moderately left to moderately right and includes a financial newspaper (*Handelsblatt*) and a conservative Swiss publication that is widely read in Germany (*NZZ*). Using the query "Reallabor", we retrieved the articles from *Neue Zürcher Zeitung*, *Die Tageszeitung* and *Die Welt* from the academic database Nexis Uni and downloaded the articles from *Handelsblatt*, *Frankfurter Allgemeine Zeitung* and *Süddeutsche Zeitung* from the respective newspaper databases. However, this first search yielded only a small number of articles. Thus we extended our search in two dimensions: we added local newspapers that are published in cities with real-world lab activities and we included keywords that are semantically related to "Reallabor". The local newspapers included *Aachener Zeitung* (North Rhine-Westphalia), *Darmstädter Echo* (Hesse), *Hamburger Abendblatt* (Hamburg), *Lausitzer Rundschau* (Brandenburg), *Sächsische Zeitung Stammausgabe Dresden* (Saxony), and *Stuttgarter Zeitung* (Baden-Württemberg). We downloaded the articles from *Hamburger Abendblatt* from the newspaper database and retrieved the articles of all other local newspapers from the academic database Nexis Uni. Table 1 gives the list of terms related to "Reallabor" that served as keywords for the extended search. Figure 1 displays the resulting number of articles by year and newspaper.

The highest number of articles were published in 2021, followed by 2020. The German government represented by the Federal Ministry for Economic Affairs and Climate Action promoted real-world labs through several initiatives that explain the increased attention to this topic in the general public. Table 2 provides an overview of the respective milestones.

We noticed that some articles that were published in a nation-wide newspaper were reprinted in one or two local newspapers. We decided to keep the second and third copy of these articles in order to be able to track the number of articles that were published in local newspapers. We found that 54.8% (613 from 1,139) of the articles in the Real-World Lab corpus were published in the six local newspapers.

Table 1
Search Terms Related to "Reallabor" that were used in the German newspaper texts

Language	Terms
German	Techniktest, Feldexperiment
English	Regulatory Sandbox, Transition Lab, Urban Lab, Future Lab, Innovation Lab, Living Lab, Social-Design Lab, Real-World Lab

Table 2
Real-World Lab Milestones Initiated by the German Federal Ministry for Economic Affairs and Climate Action

Date	Milestone
December 2018	Launch of the Strategy for Real-World Labs (or Regulatory Sandboxes)
February 2019	Real-World Lab 2019 Innovation Award Competition
July 2019	Presentation of the Innovation Award
June 2021	Publication of a revised funding concept for "Real-World Labs for the Energy Transition"
September 2021	Publication of a concept for a new Real-World Lab Law
November 2021	Real-World Lab 2022 Innovation Award Competition
May 2022	Presentation of the Innovation Award

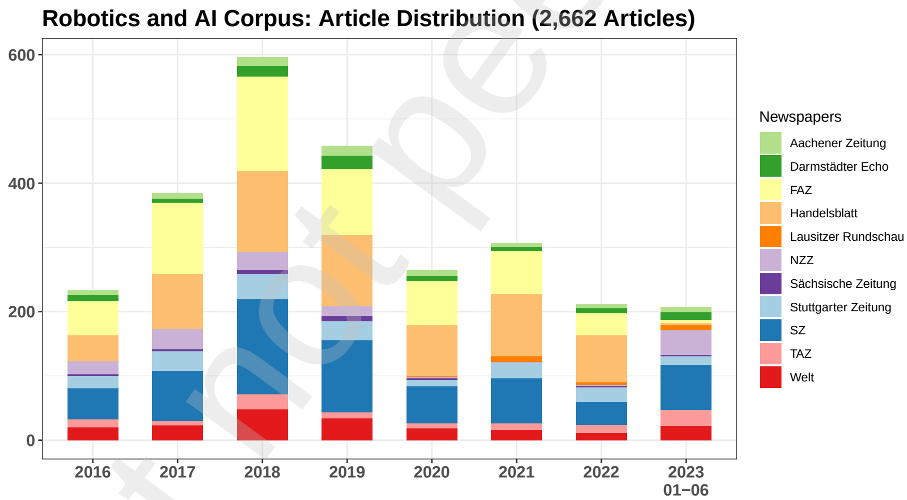


Figure 2: The distribution of 2,662 newspaper articles published in twelve German newspapers from January 1, 2016 to June 30, 2023 that feature the keywords *r/Robot** and *k/Künstlich* Intelligenz*.

For the Robotics/AI corpus, we used the query "Robot* OR robot* OR künstliche Intelligenz OR Künstliche Intelligenz" and retrieved articles from the same newspapers like for the Real-World Lab corpus. The search yielded 8,375 articles. We noticed that many articles mentioned robotics or AI only in passing and some were duplicates. We removed duplicates and kept only articles that contained both terms, "k/Künstliche Intelligenz" and "r/Robot*, and at least one of the terms more than once, resulting in a data set with 2,662 articles. Figure 2 shows the number of articles by year and newspaper.

Notably, the highest number of articles were published in 2018 and 2019 as the German government launched the National AI Strategy for Germany in 2018 and Germany and France signed a joint AI roadmap in 2019. It is also noteworthy, that the number of articles published in the first six months of 2023 is only slightly lower than the number

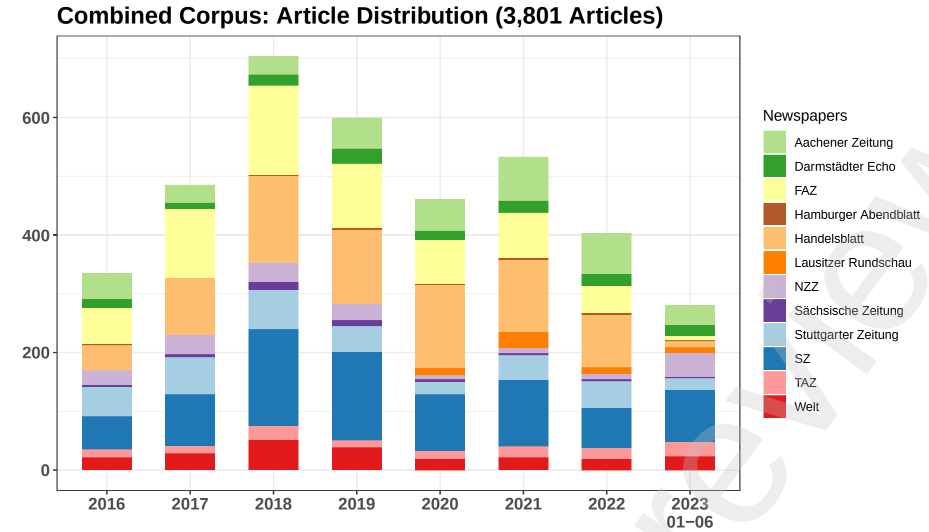


Figure 3: The distribution of 3,801 newspaper articles published in twelve German newspapers from January 1, 2016 to June 30, 2023 that feature the keywords *r/Robot** and *k/Künstlich* Intelligenz* or *Reallabor* or one of the related terms.

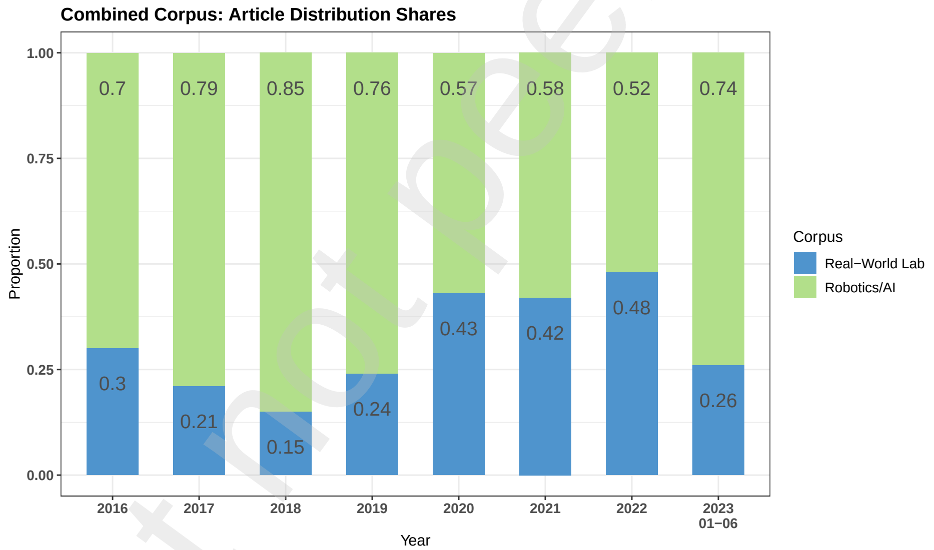


Figure 4: The articles in the combined corpus originate from two separate searches that resulted in the Robotics/AI and the Real-World-Lab corpora. Shares of the original corpora in the combined corpus by year.

of articles published in the twelve months of 2022. With the release of ChatGPT end of November 2022, generative AI as implemented in large language models attracted the attention of the general public and was widely reported on in 2023.

For this study we combined the two corpora as we are interested to uncover the connections between robotics and AI and real-world labs. Figure 3 shows the distribution of the articles in the combined corpus by year and newspaper.[5] Figure 4 provides information about the sub-corpus shares in the combined corpus.

[5]Note that there was no overlap between the two sub-corpora.

3.3. Preprocessing

Raw text data is routinely preprocessed before it is used as input in a model. We carried out the following procedure: First, we applied lemmatization. Lemmatization is the process of reducing a word to its canonical form. German is a richly inflected language and many words have the same canonical form. For example, the lemma (canonical form) of all the following words is ”groß”: ”größer”, ”größte”, ”größeres”, ”größten”. We performed lemmatization manually. Second, we included a small number of bigrams. The unit of analysis in this study is a single term. However, given the subject matter, we included the following bigrams to preserve their meaning: „Vereinigte Staaten”, „Silicon Valley”, ”Wall Street”, and ”Science Fiction”. We concatenated the two terms to create a single term with camel case. For example, we mapped ”Vereinigte Staaten” to ”VereinigteStaaten”. Third, we removed punctuation marks. We wished to preserve the names of German member states, so we removed the hyphen and concatenated the two words, generating new compound words written in camel case. For example, ”Baden-Württemberg” became “BadenWürttemberg”. Fourth, we removed numbers. Fifth, we removed stop words. Stop words are words that appear frequently in natural language, but carry little meaning for the purposes of our analysis. They include articles (”der”, ”die”, „das”), pronouns (”er”, ”sie”, ”es”), prepositions (”unter”, ”über”) and first names. We used a standard stop word list for the German language and extended it by a custom stop word list[6]. We included the search terms in the custom stop word list for better results. Sixth, we removed short words with less than four letters. Seventh, we dropped terms that featured less than 30 times in the corpus. The contribution of these words to the topics in the model is negligible, the algorithm becomes faster without loss of statistical information. Note that we decided against transforming all words to lower case, since we wanted to preserve nouns, which are spelled with capital letters in German. Words that are capitalized since they appear in the first position of a sentence are mapped to their canonical forms through lemmatization.

After preprocessing, a vocabulary was created with all terms and their frequencies in the corpus. Each document was represented as a vector of terms and their respective frequencies. The vectors were combined to a matrix, the document-term-matrix. Thus the preprocessing steps facilitated mapping the articles to a mathematical representation. Figure 9 in Appendix B illustrates the preprocessing process.

3.4. Model Specification

We chose to leverage the Structural Topic Model (STM) to evaluate the content of our newspaper article collection and to uncover hidden structures in the content. We used the R package stm, an implementation of the STM algorithm (Roberts, Stewart and Tingley (2019)). The input of the model is the document-term matrix and optionally, metadata in the form of covariates. The output of the model are topics and estimates about the relationships between the topics and the included covariates. Topics are probability distributions over all terms in the vocabulary and documents are probability distributions over all topics. Thus each topic is a combination of terms with varying proportions and each document is a combination of topics with varying proportions.

The model is specified by the number of topics k and the definition of covariates. The number of topics k must be chosen carefully. We determined k iteratively: we generated two sets of candidate models and selected the model with the best topic quality. Topic quality is a combination of semantic coherence and exclusivity of words to topics (Roberts et al. (2014); Roberts et al. (2019)). Semantic coherence measures the degree to which the content of a topic is meaningful. In a topic with high semantic coherence the most widely used words frequently occur together. Exclusivity measures the uniqueness of the terms in a topic compared to the terms in the other topics. A topic with high exclusivity has many terms that are unique to this topic. For our analysis, we chose the model with $k = 32$, as it performed best in terms of semantic coherence and exclusivity[7].

We included two covariates: an indicator variable showing whether the article belongs to the Robotics/AI or the Real-World Lab corpus and the publication date. The indicator variable is used to model the prevalence contrast of the topics given the two underlying corpora. The topical prevalence contrast is a measure that determines whether a topic is categorized as a Robotics/AI topic, an integrated topic or a Real-World Lab topic. The integrated topics are of particular interest for this study.

Based on the publication date, for each day in the time range, the algorithm estimates the shares of every topic using the least squares method on a polynomial of degree ten. Note that the shares of all topics add up to 1 for each day. Including this covariate facilitates an analysis of topic dynamics that traces the change of topic shares over time. Topic shares represent the relative importance of the topics at a specific time, where prominent topics have high shares. The

[6] To illustrate, we listed some example words from the custom stop word list in Appendix A.

[7] For details on the iterative process for selecting the value of $k = 32$, see Appendix D.

fluctuation of topic shares reveals the topics that shape or dominate the discourse at a particular point in time. Figure 10 in Appendix C illustrates the modeling process.

4. Results

4.1. Topic Labeling, Top Terms and Topic Proportion

We surveyed the 32 topics in our model and labeled them manually. We chose the labels after closely reading the five most important articles in each topic[8] and inspecting the respective word clouds[9].

Table 3 provides an overview of our 32 topics, the six most frequent words per topic, and the topic proportions disregarding the timeline[10]. Note that the topic numbers are an output of the algorithm and may serve as a short reference for the topics. The top three topics are T3, labeled ”Machine Learning and AI Development Methods”, T30, labeled ”Philosophical Considerations on the Digital Transformation”, and T5, labeled ”Digitization of Business Processes” with overall proportions of 5.87%, 5.17% and 4.68% respectively. T3 and T5 are Robotics/AI topics, which is expected, given that the share of the Robotics/AI corpus is 70% of the combined corpus. T30 is a topic that draws heavily on both corpora and is therefore considered a bridge topic. T28, labeled ”Real-World Labs for Ecofriendly Mobility in Aachen” is the Real-World Lab topic with the overall highest proportion (4.15%) in the combined corpus.

Broader themes include ”Robots and AI Applications in Professional Contexts” (T5, T6, T8, T11, T17, T20, T22, T29, T31, T32), "Government and Private Funding" (Topics T1, T7, T10, T16, T25, T27), ”Real-World Labs” (T9, T15, T18, T28), ”The Work Place" (T2, T26), ”Robots and AI in Private Households and Schools” (T19, T23), ”Robots and AI in Movies, Art and Literature” (T4, T21), and ”Theory” (T3, T30). These are the main themes in the German media discourse about robotics, AI and real-world labs.

4.2. Topic Dynamics

In addition to topic content, the results of the modeling exercise are estimates of the relationships between the topics and the covariates. The first covariate in our structural topic model is the publication date of the articles. Including this covariate facilitates an analysis that traces changes of topic shares over time. Topic shares can be seen as proxy for the relative importance of the topics at a specific time, where prominent topics have high shares. The change of topic shares identifies the topics that shape or dominate the discourse at a particular point in time. Figure 5 displays the shares of the topics with the highest shares at some point in time between January 1, 2016 and June 30, 2023.[11]

From the start of 2016 to May 2019 the topic ”Machine Learning and AI Development Methods” (T3) was the dominant topic. Disregarding the timeline, this topic has the highest overall share (5.87%)[12] and is based on articles on deep learning, artificial neuronal networks, evolutionary computation and the development of a general AI. It contains reporting on the technological progress in the development of AI. The topic ”Artificial Humans in Movies and Literature” (T4) is the fourth most important topic overall and its shares vacillated depending on new movies being released and new novels being published and reviewed in the newspapers. It was the prominent topic for several months in 2019. The topic ”Real-World Labs for the Energy Transition” (T9) began gaining attention mid 2018 and from the start of 2020 to October 2022 this topic remained one of the two top topics. From May 2020 to March 2021 the topic ”Covid Mass Vaccination as a Field Experiment” (T13) superseded the topic about energy real-world labs as the top topic. The shares of topic ”Chatbots and ChatGP” (T6) increased rapidly towards the end of the evaluation period and became the dominant topic in October 2022. We will discuss the results of this analysis in section 4.6.

4.3. Topical Prevalence Contrast

The second covariate is an indicator variable showing whether the article belongs to the Robotics/AI or to the Real-World Lab corpus. This variable is used to elicit a topical prevalence contrast, a measure of the variability of topic coverage conditional on the sub-corpus. The results of our topical prevalence contrast analysis are given in Figure 6.

The topics on the left-hand side draw mostly from articles in the Robotics/AI corpus. The topics that are most notably based on the Robotics/AI corpus include ”Machine Learning and AI Development Methods” (T3), ”Artificial Humans in Movies and Literature” (T4) and ”Automation Consequences for the Workplace” (T26). The topics on the

[8] Note that each article is assumed to be a distribution of topics with various shares s. Given a topic T, the most important articles for T are those with the highest values for s_T.

[9] The word clouds of selected topics are displayed in Appendix E.

[10] A table with the topics ordered by Topic Number is given in Table 4 in the Appendix F.

[11] We omitted T15 and T28 from this graph, since these two topics are specific to Aachen and do not pertain to the whole country.

[12] Compare Table 3

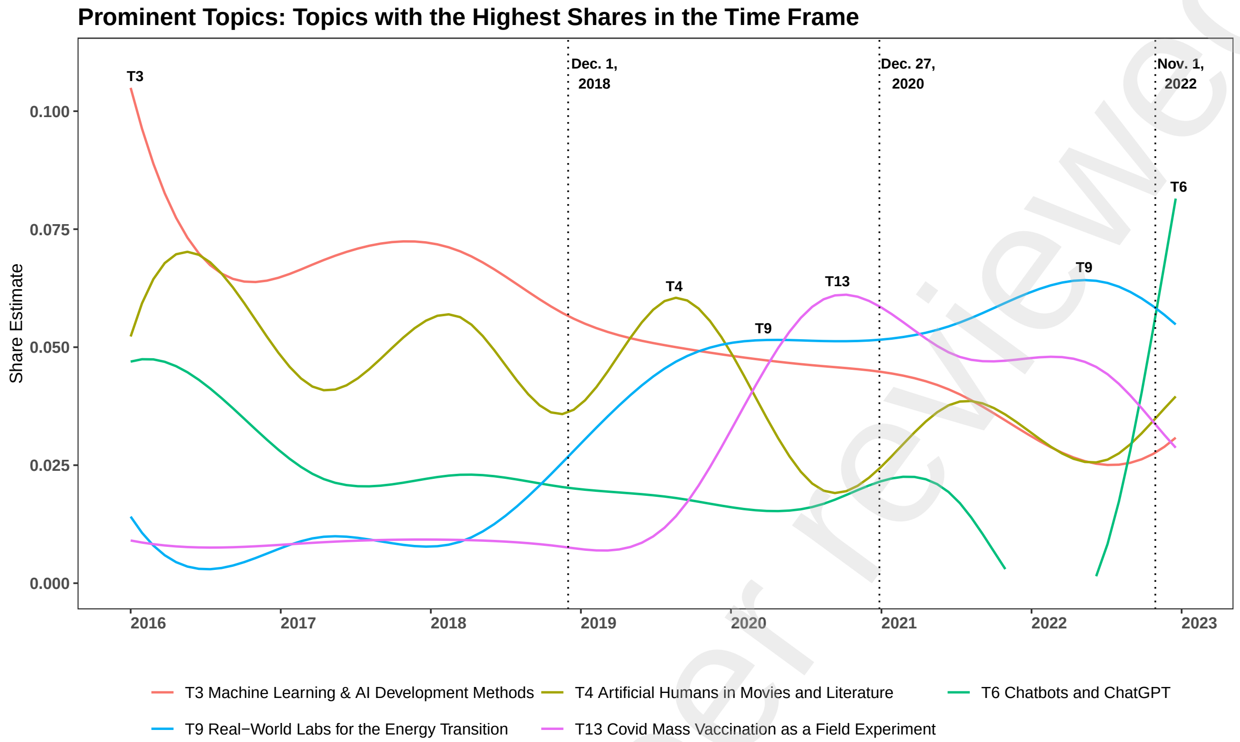


Figure 5: Topic proportions change over time. This figure shows the shares of the topics with the highest shares at some point in the time frame. Topics with the highest shares dominate the reporting. The vertical dotted lines denote three important dates: Dec 1, 2018 (the launch of the Strategy for Real-World Labs by the German Federal Ministry of Economic Affairs and Climate Action), Dec. 27, 2020 (the official start of the Covid vaccination campaign in Germany), and Nov. 1, 2022 (the month in which ChatGPT was first released). Note that T3, T4, and T6 are robotics/AI topics, T9 is a real-world lab topic and T3 is a bridge topic. Across time we see a shift of the dominating topic from RAI to RWL.

right-hand side are mainly based on articles in the Real-World Lab corpus. The three topics that are mostly based on the Real-World Lab corpus include "Real-World Labs for Ecofriendly Mobility in Aachen" (T28), "Real-World Labs for the Energy Transition" (T9), and "Future Lab Aachen - the Whole City as a Future Lab" (T15). All of these topics belong to the real-world lab theme. Note that these topics show a greater deviation from zero than the robotics/AI topics, indicating that they draw to a lesser extent from articles in the Robotics/AI corpus than vice versa.

We are particularly interested in the topics that are equally based on both sub-corpora. They are the bridge topics that link the domains of robotics/AI and real-world labs and are marked green in Figures 6 and 7. The bridge topics at the interface of robotics/AI and real-world labs include "Service Robots and Assistance Systems" (T20), "Philosophical Considerations on the Digital Transformation" (T30), "Fintech and Legal Tech" (T32), "Digitization, Robots and AI in Schools" (T23), "Tech Start-Ups" (T16), "AI and Robotics Cutting-Edge Research Funding in Bavaria" (T7), "Government Funding for Research on AI and Climate Change" (T1), "Digital Transformation Leaders" (T24), "Covid Mass Vaccination as a Field Experiment" (T13), "Robotics and AI in Construction, Agriculture and Policing" (T8), and "Fair Pay Innovation Lab"[13] (T2).

4.4. Correlation Network

In addition to topics and their proportions, the STM algorithm outputs information on the structure of the relationships between the topics. The correlation network offers a visual representation of this relationship structure. The vertices denote topics and the edges represent positive correlations, that indicate that the two topics are likely

[13]Based in Berlin, the Fair Pay Innovation Lab is an advocacy group for fair pay and equal opportunity at the work place.

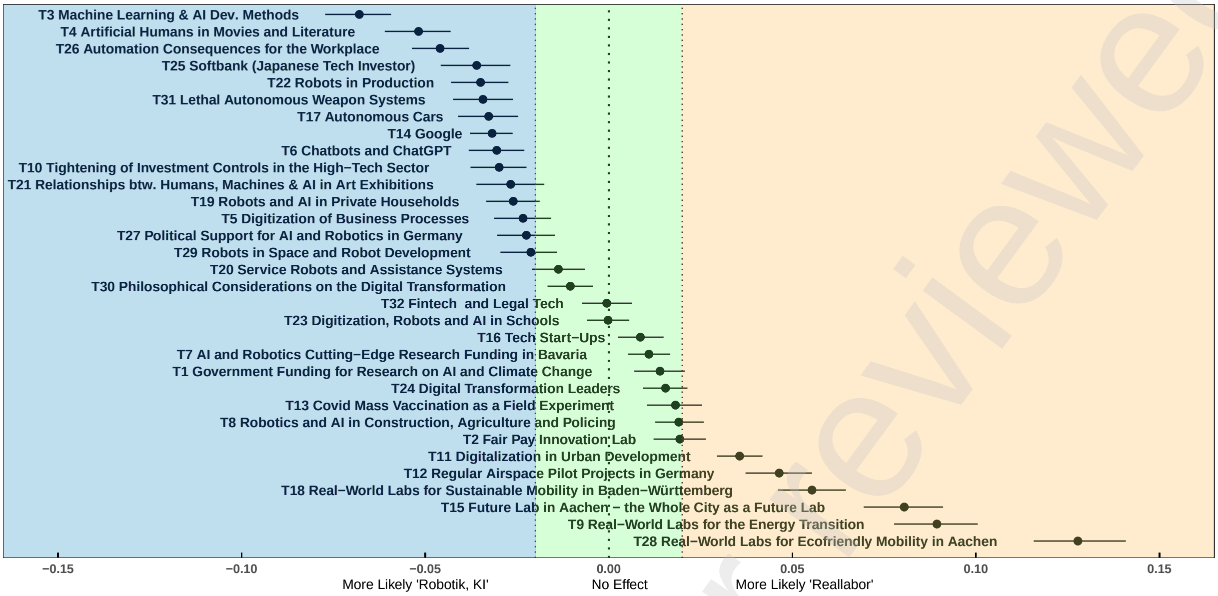


Figure 6: Topical prevalence contrast reflecting the variability of topic coverage conditional on the sub-corpus. The dots denote the means and the lines denote the 95% confidence intervals of the estimates. Topics in the blue area are predominantly based on the Robotics/AI corpus, topics in the yellow area are mostly based on the Real-World Lab corpus and topics in the green area are bridge topics that link the two domains.

discussed within the same articles. The correlation network of our model is given in Figure 7. Note that the size of the vertices represent the proportions of the topics disregarding the timeline.

We colored the vertices according to the three types of topics identified by the topics prevalence contrast analysis in the section above: blue vertices represent topics that are predominantly based on articles in the Robotics/AI corpus, yellow vertices represent topics that are mainly based on articles in the Real-World Lab corpus and green vertices represent bridge topics that link the two domains.

We observe by visual inspection that the correlation network has two main clusters that neatly divide the 32 topics into real-world labs topics (marked yellow) and non-real-world lab topics (marked blue and green). There are some edges between the two clusters.[14] Notably, these edges connect vertices that represent real-world lab topics with vertices that represent bridge topics, specifically the application topic ”Robotics and AI in Construction, Agriculture and Policing” (T8) and the three funding topics ”Government Funding for Research on AI and Climate Change” (T1), ”Tech Start-Ups” (T6), and ”AI and Cutting-Edge Research Funding in Bavaria” (T7). Three of these four topics (T1, T7, T8) are located at the margin of the non-real-world lab cluster and have the smallest distance to the real-world lab cluster.

Community detection can be used to detect topics with similar properties and extract sub-groups based on various specifications. We have tried various cluster (or community) detection algorithms that are commonly used in network analysis.[15] We analyzed the resulting clusters for several of threshold values for correlation and saw some similarities to our approach of the bridge topics. E.g. T1 acts as a link between a variety of clusters for several threshold values.

Regarding the importance or the ”prestige” of a topic within a network, Eigenvector centrality is an indicator that is frequently applied. Eigenvector centrality is an important concept in graph theory to measure the influence of a node in a connected network. Connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. A high eigenvector score thus means that a node is connected to many nodes who themselves have high scores. Sometimes, eigenvector centrality is also used to measure the ’prestige’ of a node

[14] Note that the appearance of an edge is conditional on the chosen threshold value which we have set to 0.01. If we reduced (increased) the threshold, more (less) edges appeared.

[15] We used the Louvain algorithm of the igraph library of R.

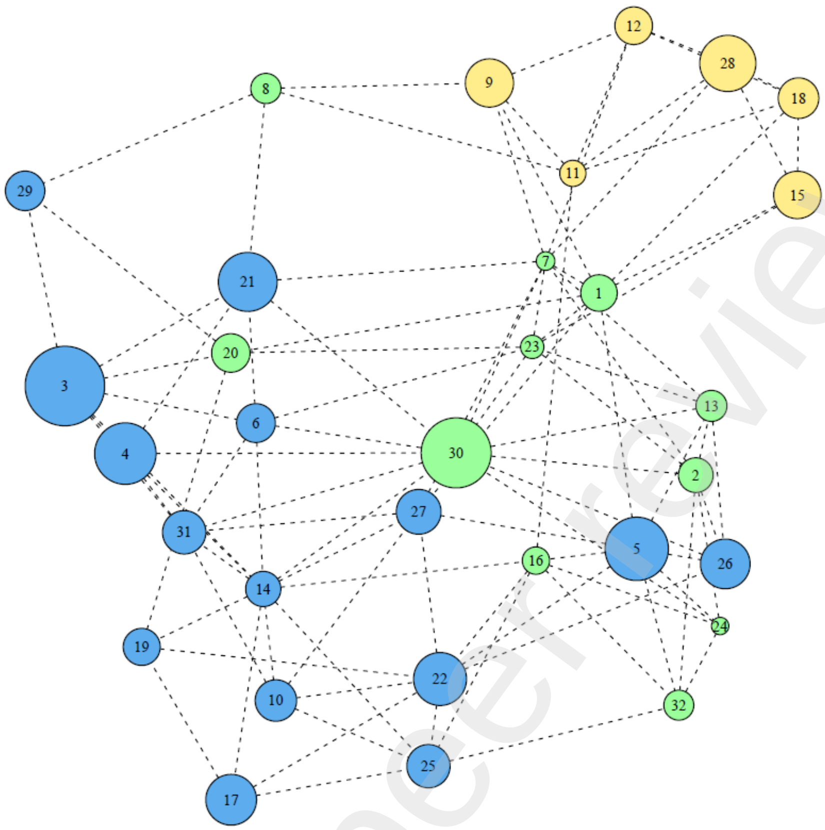


Figure 7: Correlation Network. The edges display a positive correlation between two topics and indicate that the two topics are likely to be discussed in the same articles. The threshold value for drawing edges is 0.01. The size of the vertices represent the proportions of the topics disregarding the timeline. Blue vertices represent topics based on the Robotics/AI corpus, yellow vertices represent topics mainly based the Real-World Lab corpus and green vertices represent bridge topics that link the two domains.

in a network. Interestingly, the topics with the highest Eigenvector centrality are "Philosophical Considerations" (T30) and "Political support for AI in Germany" (T27) (compare Table 4 in the Appendix).

4.5. Sentiment Analysis Based on Four Psycholinguistic Attributes

Another avenue of text analysis is the study of the emotional content of a text. Usually, only one dimension is covered. In contrast, our sentiment analysis is based on a specialist dictionary for the German language which provides values ranging from 0 to 10 for four psycholinguistic attributes for roughly 340,000 German lemmas (compare Köper and Schulte im Walde (2016)). The attributes include valence, arousal, abstractness and imageability.

Valence refers to the pleasantness of a word (unpleasant vs. pleasant), arousal rates the intensity of emotional activation inherent in a word (calm vs. alert), abstractness determines the level of sensory perceptibility (abstract vs. concrete), and imageability describes the visibility of the meaning of a word (invisible vs. visible).

We define the topic score for these four attributes as follows:

$$s_a = \sum_{w \in V} \gamma_{w,t}\, \sigma_{w,a}.$$

Here V refers to the vocabulary of the entire corpus, $\gamma_{w,t}$ denotes the estimated frequency of a word w in a topic t, and $\sigma_{w,a}$ represents the score for attribute a of word w in the dictionary. The topic scores are given in Table 4 in the Appendix F.

The articles in the corpus report on both, abstract topics (e.g. T30) and more concrete topics (e.g. T20). We address the psycholinguistic dimensions of abstractness and imageability in section 4.6.

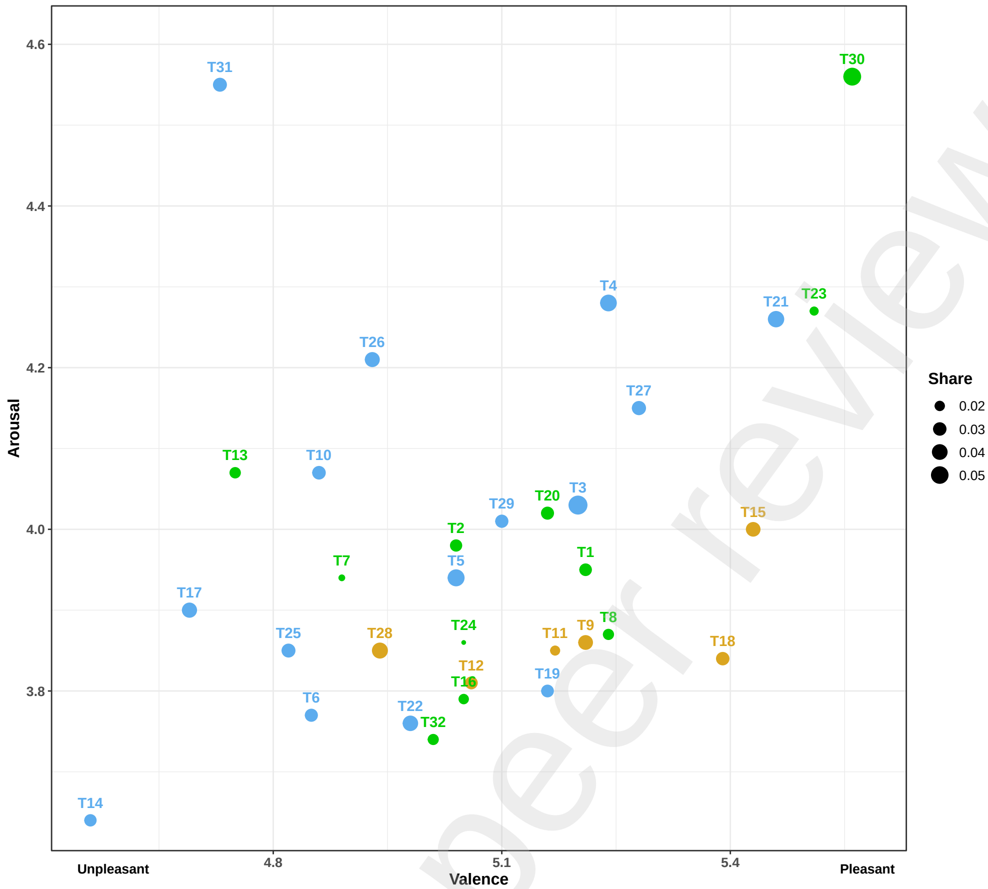


Figure 8: Topic Valence and Arousal. The size of the vertices represent the proportions of the topics disregarding the timeline. Blue vertices represent topics based on the Robotics/AI corpus, yellow vertices represent topics mainly based the Real-World Lab corpus and green vertices represent bridge topics that link the two domains.

We studied the attributes valence and arousal in more detail. Figure 8 displays the topic scores of these two attributes. Like in section 4.4 above, the colors represent the topic types (Robotics/AI topics vs. bridge topics vs. Real-World Lab topics), and the size of the dots denote the topic shares disregarding the timeline as reported by their proportions in Table 3.

As expected, the topic about lethal autonomous weapon systems (T31) has a low valence score and one of the highest arousal scores. This topic is based on articles reporting on progress in the development of lethal robots and AI that activates high levels of emotions and is generally unpleasant. Its opposite is T30 ("Philosophical Considerations on the Digital Transformation") that has the highest scores for both, valence and arousal, and is based on articles discussing the potential of the digital transformation, with high frequencies for words that evoke hope in a better future.

Other topics with high valence scores include T23 ("Digitization, Robots and AI in Schools"), T21 ("Relationships between Humans, Machines and AI in Art Exhibitions"), T15 ("Future Lab Aachen - the Whole City as a Future Lab"), and T18 ("Real-World Labs for Sustainable Mobility in Baden-Württemberg"). These are the most pleasant topics. T26 ("Digitalization and Automation Consequences for the Workplace") ranks second regarding arousal.

The most unpleasant topics on the other end of the spectrum, include T17 ("Autonomous Cars"), T13 ("Covid Mass Vaccination as a Field Experiment"), and T14 ("Google"). T17 draws from reporting on the dangers of autonomous mobility and T13 is based on discussions of the high mortality rates of the Covid pandemic, so the low valence scores of these topics are easily comprehensible. However, the low valence score of T14 comes as a surprise and needs further investigation.

4.6. Discussion

In section 4.2 we have shown how the focus of the reporting on RAI and RWL in German broadsheet newspapers has changed over time. For more than three years from the start of our time frame, the topic ”Machine Learning and AI Development Methods” (T3) was most reported on. It is remarkable that this topic held such prominence for such a long time, given that the readers of the newspaper articles were not experts in the field, but the general public. This demonstrates that the German public was intrigued by technical details of machine learning and AI development and attentive to emerging developments.

The topic ”Artificial Humans in Movies and Literature” (T4) was one of the top two topics for the first four years of our time frame. In this period it was the most important topic next to the topic about machine learning and AI development methods (T3). While T3 kept the public abreast of technical details, T4 captured and nurtured their imagination, demonstrating the implications of living with advanced technologies through fictional stories. The comparatively high valence and arousal values (see Figure 8) indicate that these stories were mostly inspiring and hopeful.

The topic ”Covid Mass Vaccination as a Field Experiment” (T13) was prominent for ten months in 2020 and 2021, the period when strategies to contain the Covid-19 pandemic were hotly debated and mass vaccinations against the virus were perceived as a field experiment. This topic is one of the bridge topic that links the two domains under consideration and it stands out as one of the topics with the lowest valence values (see Figure 8), easily explainable on the grounds that Covid had initially such high mortality rates.

The topic ”Real-World Labs for the Energy Transition” (Topic 9) began gaining attention mid 2018 and from the start of 2020 to October 2022 this topic remained one of the two top topics. This ranking reflects the significance of the real-world labs for the energy transition, a new project type established by the Federal Ministry for Economic Affairs with the publication of the 7th Program for Energy Research in September 2018. These guidelines for energy research policy introduced real-world labs for the energy transition as test spaces for innovation and regulation, designed to accelerate the maturation process of innovative energy technologies. Table 2 lists the milestones in the development of this new project and funding format.

Finally, the shares of topic ”Chatbots and ChatGP” (Topic 6) increased rapidly towards the end of the evaluation period and became the dominant topic in October 2022. The reason for this astonishing rise is the release of the generative AI ChatGPT in November 2022, the media attention it garnered and the hype it entailed. However, this topic has a low valence value (see Figure 8), indicating that it was predominantly discussed in negative terms, though unemotionally, as suggested by its low arousal value.

In section 4.3 we have identified the topics that are evenly based on both sub-corpora and have introduced the term ”bridge topics” for them, as they link the domains of robotics/AI and real-word labs, as depicted in Figure 6. Note that the two bridge topics with strong links to research activities, ”Government Funding for Research on AI and Climate Change” (T1) and ”AI and Robotics Cutting-Edge Research Funding in Bavaria” (T7), belong to the broader theme ”Public and Private Funding”. As shown in section 4.4, both topics are among the few topics from the non-real-world lab cluster that have edges to the real-world lab cluster in the correlation network (see Figure 7). We conclude that these topics are essential for linking the two domains. Further, the bridge topics ”Service Robots and Assistance Systems” (T20), and ”Robotics and AI in Construction, Agriculture and Policing” (T8) belong to the broader theme ”Robots and AI Applications in Professional Contexts”. In the correlation network, T20 also serves as a topic that links the real/world lab cluster to the non-real-world lab cluster, indicating that it is particularly suited to act as a linkage between the two domains robotics/AI and real-world labs.

We also observed that some overarching themes are discussed in very different ways. For example, there are two topics about mobility: the topic on autonomous cars (T17) is part of the RAI domain and is viewed rather critically and the topic on sustainable mobility (T18) is part of the RWL domain and viewed very positively. T19 and T20 relate to robotic assistance systems, both of which appear with high valence and somewhat different arousal scores.

Note that the network in Figure 7 does not display three distinct clusters that correspond to the three types of topics identified in Figure 6. Rather, the network reveals a structure that may be grouped to clusters in various ways and that needs further analysis which we postpone to follow-up work. However, our focus on prevalence contrast is novel and provides additional insights that could not be gained from standard network metrics.

As argued before, the topics with the highest Eigenvector centrality are ”Philosophical Considerations” (T30) and ”Political Support for AI in Germany” (T27) (compare Table 4 in the Appendix).

We also applied a correlation analysis covering the dimensions of topic proportion, Eigenvector centrality, topical contrast and the four psycholinguistic attributes disregarding the timeline. The results are shown in Figure 14 in the

Appendix G. Clearly, the size of the topic proportion has almost no impact on the other dimensions discussed (with the exception of arousal that weakly correlates with size (0.367*)).

Looking at the correlation plot in Figure 14, the intuition that more abstract and less imageable topics are less prestigious (as reflected by low Eigenvector centrality values) and thus negatively correlated with Eigenvector centrality is confirmed (-0.590** for imageability and -0.569*** for abstractness). In contrast, arousal and Eigenvector centrality are positively correlated (0.394*). The negative correlation between topical prevalence contrast and the Eigenvector centrality (-0.560***) highlights that mostly topics from the RAI subcorpus are linked to other prestigious topics[16]. Finally, topics originating more from the RWL sub-corpus trigger less arousal.

5. Critical Reflection and Policy Implications

As with all scientific analyses, a classification of the results requires a close examination of the data used. The data set consists of two sub-corpora, which are based on well-founded search terms. These were carefully determined in advance and a broader perspective was chosen in the case of the relatively new concept of real-world laboratories, which are not yet clearly defined and which include a wide range of English-language terms even though the language of the newspaper was German.

We chose big German broadsheet newspapers that cover a broad political spectrum as data sources, expecting as a result a comprehensive and undistorted picture of reporting in the fields of RAI and RWL. These newspapers were selectively supplemented by regional newspapers from cities with a variety of real-world laboratory activities. We made sure that only one regional newspaper was used per federal state (e.g. in Baden-Württemberg the *Stuttgarter Zeitung*, but not the *BNN* from Karlsruhe). Furthermore, not all federal states were covered.

As a result, we obtained a combined text corpus where the RAI sub-corpus is twice as large as the RWL sub-corpus. This imbalance seems harmless for the analysis, and we can show that over time there is a shift in the dominant topics towards RWL-related topics (cf. the topic diffusion curve for T9 in Figure 7)[17]. The analysis does not include social media, although this would be possible from a methodological point of view. Reference is made here to future work.

The results obtained and the interpretations derived from them strongly depend on the preprocessing of the data, which should only be partially automated. In particular, compiling the custom stop word list requires careful consideration[18]. One advantage of the German language is the creation of a comprehensive sentiment analysis along four dimensions of psycholinguistic attributes, which goes far beyond commonly used sentiment analyses.

The selection of the number of topics and the topic labeling are particularly critical for the outcome. While the selection of the best possible number of topics can be supported by the algorithm, this is not feasible for topic labeling. In order to obtain accurate labels, we have chosen a combined view of the top terms and the top articles from which the topic orignates. This text corpus also shows that not all topics can be conclusively assessed and consistently included in the overall context (here, for example, this applies to T14 ”Google”). In particular, as with all AI methods, one must be aware that the algorithm is not neutral.

The following strengths of the chosen approach are particularly obvious. It enables the analysis of large volumes of text that cannot be analyzed manually in this depth. The STM facilitates mapping the text data to a mathematical representation and thus to link it with other analysis methods like network analyses, eigenvector centrality and sentiment analysis based on a dictionary with psycholinguistic attributes. This combination of methods and indicators enables a comprehensive analysis of the corpus and there are a number of extensions that go far beyond what is described here.

The present approach is only a first step that could be followed by many other analyses. These include a much more comprehensive network analysis, the comparison of standard community detection with the idea of bridge topics developed here, as well as the additional consideration of the strength of topic correlation, which varies between the topics. In addition, a particular challenge is to explore the interfaces with other established indicators. Furthermore, it is possible to search for specific terms in topics and to view the text corpus from this perspective. Contentwise, such a targeted analysis is basically possible in all directions and depends very much on the chosen research question. We see great potential here.

Regarding content, our aim was to identify unifying elements in complex media reports as well as idiosyncratic topics. Such a comprehensive approach is particularly important in times of overlapping challenges.

[16]Recall, that the values for topical prevalence contrast result from the analysis in Figure 6. A positive contrast value implies that the topic originates more from the RWL sub-corpus whereas a negative contrast value implies that the topic is more closely related to the RAI sub-corpus.

[17]Other studies with comparable unbalanced datasets also show that this is harmless for the interpretation of the results, e.g. geothermal energy.

[18]As the German language includes many compound nouns, *n*grams do not play a significant role. This is different in English and other languages. Working with text data in these languages, *n*grams must be chosen carefully.

Both robotics and regulatory sandboxes are important current funding lines in Germany. Politicians often face budgetary constraints and are called to prioritize funding. Especially in times of increasingly tight budgets, it is therefore all the more important to identify links between different funding contents and to promote complementarities. In modern innovation system theory, it has long been known that, in addition to companies, society as a stakeholder also plays an important role in the development and, in particular, the diffusion of technologies.

The proposed method can be utilized to obtain an understanding of recent discussion and thus serve as information for experts who design policy instruments. We suggest that the government should focus on the connecting elements in order to exploit possible complementarities in the two fields of RAI and RWL, but also beyond, in the further promotion of innovation.

6. Conclusion

Modern robots that operate outside well-structured environments and interact with humans are among the most promising technologies for future applications. Concrete requirements for the machines are still unclear, as are the framework conditions for testing the machines. The framework needs to be set up in such a way that it allows for continuous testing, while at the same time providing legal certainty for innovators and meeting the needs of potential customers. Regulatory sandboxes are a modern tool of innovation policy to enable precisely this and to involve the state as a learning actor. But how does society view these two phenomena together? Newspapers are a traditional and long-standing medium for conveying information to the public, including scientific and technological developments, and placing them in a larger context.

Both RAI and RWL are current topics in innovation promotion, but are mostly treated in isolation. This paper uses quantitative text analysis to examine 3,800 German newspaper articles in the period 2016-2023. We are particularly interested in the interface between RAI and RWL. We show that in our combined corpus, the dominant topic has changed over time from "Machine Learning and AI Development Methods" to "Real-World Labs for the Energy Transition". The connecting themes are diverse and include philosophical and legal considerations as well as specific application areas for robots, e.g. in schools and also point to the importance of public funding.

A particular focus was placed on a quantitative analysis and the linking of these two perspectives that are otherwise considered in isolation. Fostering technology adoption is a pivotal step in the innovation process and represents a well known bottleneck for the successful deployment of a new technology from research and development to the market and end user (Grubb et al. (2021)).

Ideally, by bridging traditional and new perspectives, the role of regulatory sandboxes in innovation system research can be further substantiated, findings from regulatory sandbox research can be scaled and transferred to other contexts, and the impact of regulatory sandbox research can be better evaluated and scaled. To conclude, the potential of regulatory sandbox should be exploited to a greater extent by using it in a wide variety of contexts. Sustainability transformation and technology development at the interface with society are two examples that can be expanded. However, regulatory sandboxes need to be rigorously complemented by other - well established and already better understood - methods and theories in order to gain deeper insights and scale up their societal and scientific impact.

Smooth co-evolution of technological development and the institutional environment are essential to enhance aggregate productivity and international competitiveness and to continuously increase social welfare.

Appendix

A. Excerpt from the Custom Stop Words List for Illustration

"Roboter" , "Robotik" , "Robotics" , "Robot" , "Roboterauto" , "Industrieroboter" , "Roboterarm" , "Sexroboter" , "Robotaxi" , "Roboterhersteller" , "Pflegeroboter" , "Roboterwagen" , "Killerroboter" , "Roboterhund" , "Robotic" , "Robotertechnik" , "Serviceroboter" , "Saugroboter" , "Robotiker" , . . . , "Dabei" , "Dadurch" , "Dafür" , "Dagegen" , "Daher" , "Dahinter" , "Damals" , "Damit" , "Danach" , "Daneben" , "Daniel" , "Dann" , "Daran" , "Darauf" , "Daraus" , "Darin" , "Darum" , "Darunter" , "Darüber" , "Dass" , "David" , . . . , "voneinander" , "vorab" , "voran" , "voraus" , "voraussichtlich" , "vorbei" , "vorerst" , "vorhanden" , "vorhandenen" , "vorher" , "vorigen" , "vorn" , "vorne" , "vorrangig"

B. Preprocessing Text Data

The preprocessing steps are shown in Figure 9.

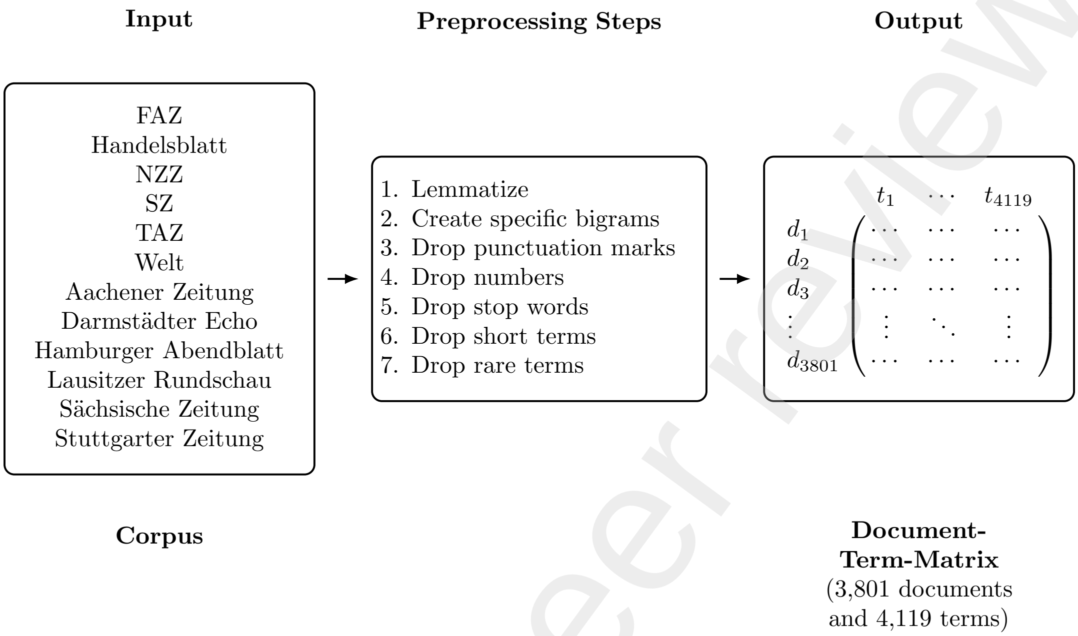


Figure 9: The preprocessing steps that take the collection of newspaper articles (corpus) and facilitate mapping it to a mathematical representation, the document-term matrix. Compare Loewe et al. (2024) for a similar presentation.

C. Model Specification

The model specification can be found in Figure 10.

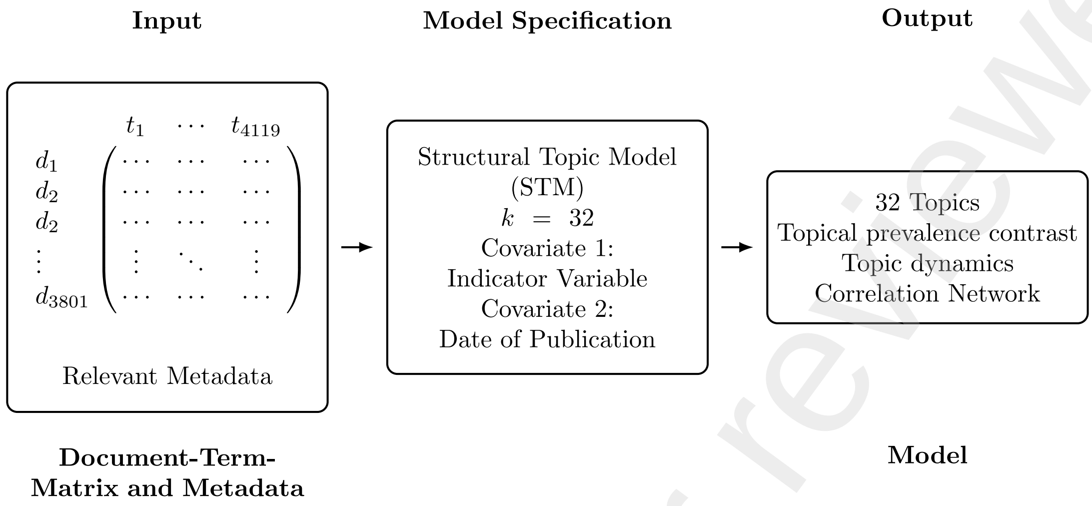


Figure 10: The Structural Topic Model (STM) is specified by the number of topics k and the definition of covariates. It takes as input the document-term matrix and relevant metadata and outputs k topics and an estimate of the relationship between the topics and the covariates. Compare Loewe et al. (2024) for a similar representation.

D. Iterative Process for Selecting the Value of k

In structural topic models (STMs), the number of topics k is given by the modeler. We selected k based on semantic coherence and exclusivity and proceeded as follows. We first generated a set of candidate models for all values $k \in [10, 60]$ in steps of 5. We computed the means of topic exclusivity and semantic coherence for each value of k and visualized the result, see Figure 11. We observed that in this set of models the model with $k = 30$ performed best.

In the second step we generated another set of candidate models with values of k in the vicinity of 30, namely $k \in [25, 40]$ in steps of 1. We compared the distribution of topic semantic coherence and exclusivity of each model, not just their mean values. We observed that the model with $k = 32$ performed best, the respective distribution is given in Figure 12.

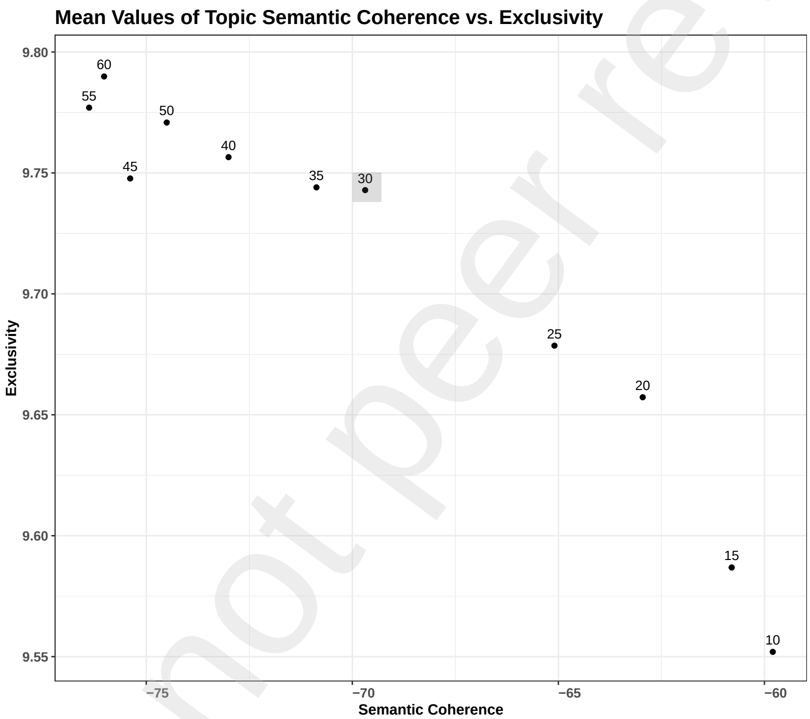


Figure 11: Finding the optimal k. Mean values of topic semantic coherence and exclusivity for the first candidate model set. The labels denote the values of k.

Table 3
Topics and their Proportions (Disregarding the Timeline), Ordered by Proportion

Topic Label	#	Top Words	Proportion
Machine Learning and AI Development Methods	3	Maschine, System, lernen, Algorithmus, Computer, menschlich	0.0587
Philosophical Considerations on the Digital Transformation	30	brauchen, wissen, Frage, Welt, Zeit, denken	0.0517
Digitization of Business Processes	5	digital, Digitalisierung, Unternehmen, Daten, Technologie, nutzen	0.0468
Artificial Humans in Movies and Literature	4	Maschine, menschlich, Buch, Leben, Welt, Menschheit	0.0454
Relationships between Humans, Machines and AI in Art Exhibitions	21	Kunst, Künstler, Film, zeigen, Ausstellung, Bild	0.0435
Real-World Labs for Ecofriendly Mobility in Aachen	28	Stadt, Aachen, Templergraben, Straße, Innenstadt, Auto	0.0415
Robots in Production	22	Unternehmen, Industrie, Bosch, Konzern, Produktion, Siemens	0.0391
Autonomous Cars	17	Auto, autonom, fahren, Fahrzeug, Tesla, selbstfahrend	0.0375
Digitalization and Automation Consequences for the Workplace	26	Arbeit, Digitalisierung, Arbeitsplatz, Maschine, Automatisierung, Beruf	0.0367
Real-World Labs for the Energy Transition	9	Wasserstoff, Energie, Projekt, Strom, Energiewende, Region	0.0356
Future Lab in Aachen - the Whole City as a Future Lab	15	Aachen, RWTH, Future, Stadt, Projekt, Veranstaltung	0.0350
Political Support for AI and Robotics in Germany	27	Deutschland, Europa, deutsch, Forschung, Wirtschaft, Industrie	0.0331
Lethal Autonomous Weapon Systems	31	autonom, Russland, System, ethisch, Waffe, Entscheidung	0.0321
Softbank (Japanese Tech Investor)	25	Japan, Dollar, Unternehmen, Softbank, Toyota, Fond	0.0320
Tightening of Investment Controls in the High-Tech Sector (USA, EU, Germany)	10	China, Land, Unternehmen, Europa, Amerika, Staat	0.0306
Real-World Labs for Sustainable Mobility in Baden-Württemberg	18	Stuttgart, Projekt, Baden, Campus, Idee, Württemberg	0.0300
Robots in Space and Robot Development	29	entwickeln, Forscher, arbeiten, Forschung, System, Entwicklung	0.0292
Chatbots and ChatGPT	6	Text, Chatbot, ChatGPT, Nutzer, Software, schreiben	0.0287
Service Robots and Assistance Systems	20	Darmstadt, Pepper, Einsatz, Patient, Pflege, helfen	0.0284
Regular Airspace Pilot Projects in Germany	12	Hamburg, Bahn, Berlin, Drohne, deutsch, Projekt	0.0278
Robots and AI in Private Households	19	Gerät, Smartphone, Amazon, Nutzer, Alexa, Produkt	0.0274
Government Funding for Research on AI and Climate Change	1	Hochschule, Universität, Professor, Studierender, Dresden, Forschung	0.0270
Google	14	Google, Amazon, Unternehmen, Konzern, Facebook, SiliconValley	0.0260
Fair Pay Innovation Lab	2	Mitarbeiter, Unternehmen, Kollege, arbeiten, wenig, Gehalt	0.0256
Covid Mass Vaccination as a Field Experiment	13	Pandemie, Corona, Krise, zeigen, Studie, wenig	0.0229
Robotics and AI in Construction, Agriculture and Policing	8	Landwirtschaft, Labor, Feld, Tier, wenig, Pflanze	0.0223
Fintech and Legal Tech	32	Kunde, Bank, Produkt, Blockchain, digital, Branche	0.0222
Tech Start-Ups	16	Start, München, Gründer, Firma, Unternehmen, Investor	0.0202
Digitalization in Urban Development	11	Stadt, smart, City, Zukunft, urban, Wohnung	0.0193
Digitization, Robots and AI in Schools	23	Kind, Schule, Schüler, lernen, studieren, digital	0.0172
AI and Robotics Cutting-Edge Research Funding in Bavaria	7	Bayern, Corona, Söder, München, Woche, gelten	0.0138
Digital Transformation Leaders	24	Unternehmen, Deutschland, gründen, Konzern, digital, Innovation	0.0128

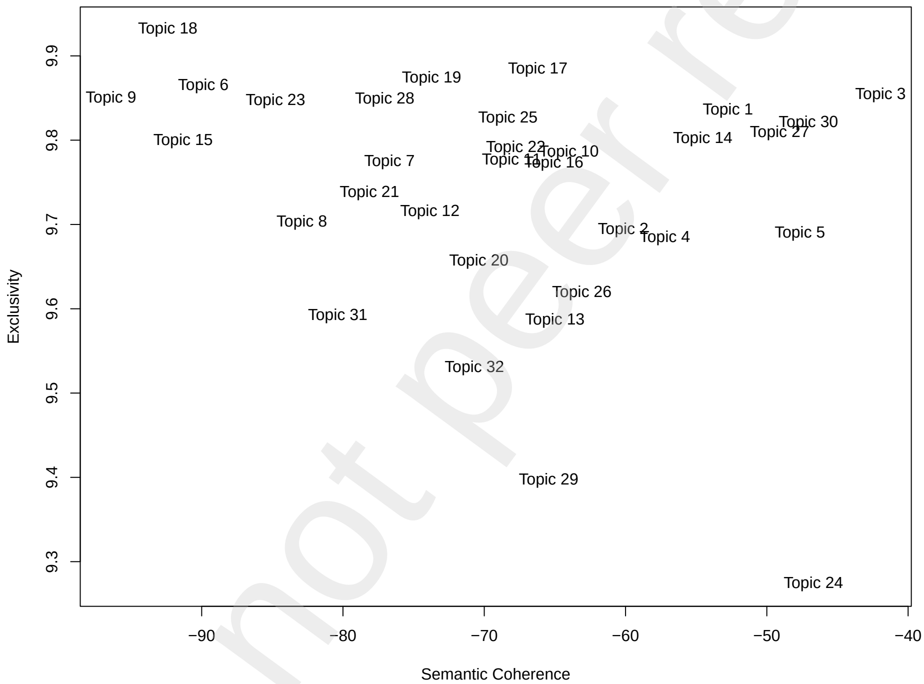
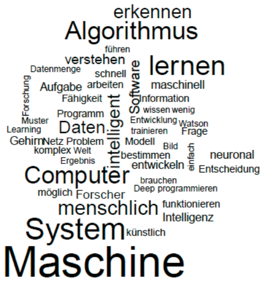


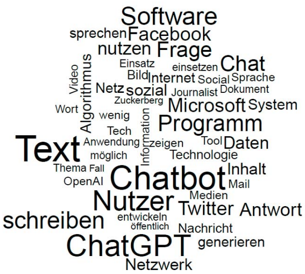
Figure 12: The distribution of the values for semantic coherence and exclusivity for all topics in the model with $k = 32$.

E. Word Clouds for Selected Topics: Top 50 Words



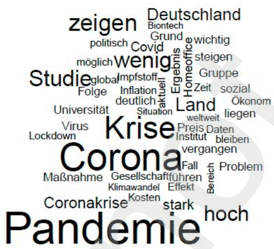
(a) Machine Learning and AI Development Methods (Topic 3)

(b) Artificial Humans in Movies and Literature (Topic 4)



(c) Chatbots and ChatGPT (Topic 6)

(d) Real-World Labs for the Energy Transition (Topic 9)



(e) Chatbots and ChatGPT (Topic 13)

Figure 13: Word clouds for topics that were dominating the discourse (compare Figure 5). These topics are discussed in more detail in sections 4.2 and 4.6. The size of the words is an indicator of their proportions (shares) in the topic. Recall that a topic is defined as a distribution of words.

F. Topic Scores

Table 4 lists the topic scores for the four psycholinguistic attributes and Eigenvector centrality.

G. Correlations Between Four Psycholinguistic Attributes, Topical Prevalence Contrast, and Eigenvector Centrality

Figure 14 highlights the correlations between four psycholinguistic attributes, topical prevalence contrast, and Eigenvector centrality.

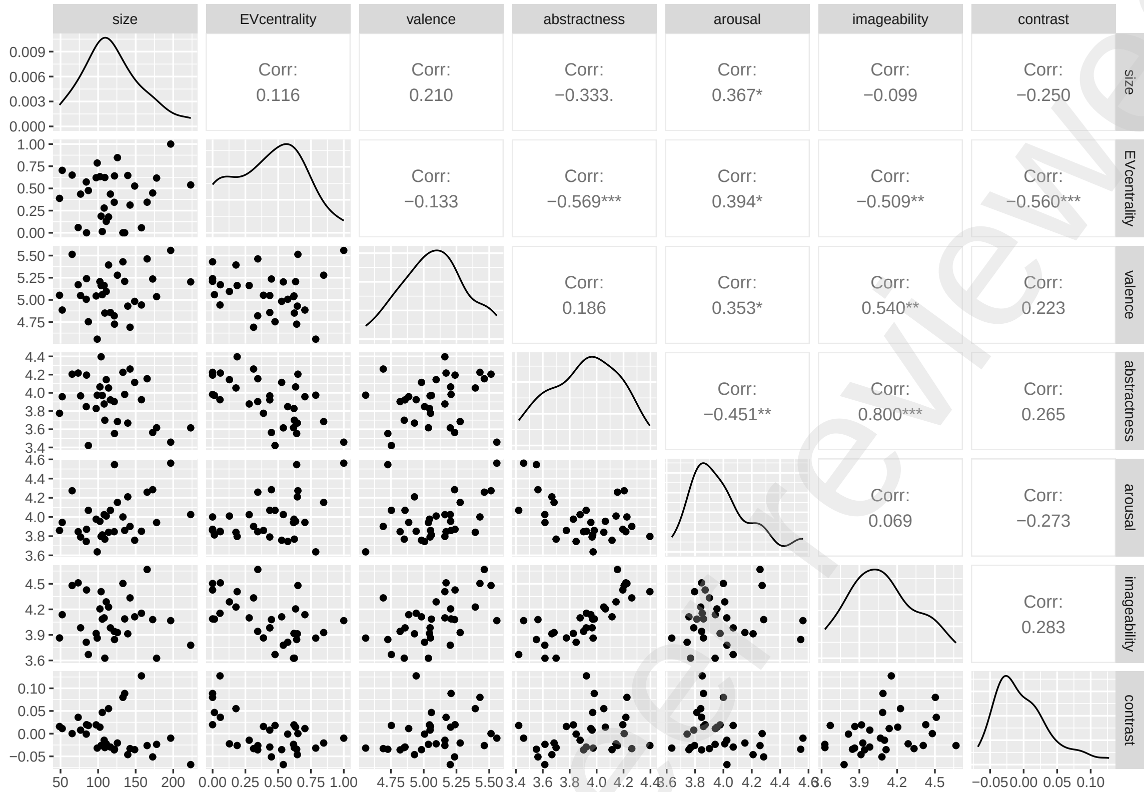


Figure 14: Correlations between key variables; own calculation based on the full text corpus and disregarding timeline.

H. Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used DeepL and DeepL Write in order to to improve the linguistic quality of the article. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Agrawal, R., Wankhede, V.A., Kumar, A., Luthra, S., Majumdar, A., Kazancoglu, Y., 2022. An exploratory state-of-the-art review of artificial intelligence applications in circular economy using structural topic modeling. Operations Management Research 15, 609—-626.

Bekar, C., Carlaw, K., Lipsey, R., 2018. General purpose technologies in theory, application and controversy: a review. Journal of Evolutionary Economics 28, 1005–1033. doi:10.1007/s00191-017-0546-0.

Blei, D.M., Lafferty, J.D., 2007. A correlated topic model of science. The Annals of Applied Statistics 1, 17–35. doi:10.1214/07-aoas114.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. Journal of Machine Learning Research 3, 993–1022.

Bresnahan, T.F., Trajtenberg, M., 1995. General purpose technologies: Engines of growth? Journal of Econometrics 65, 83–108.

Cage, J., Hengel, M., Hervé, N., Urvoy, C., 2024. Hosting Media Bias: Evidence From the Universe of French Broadcasts. 2002-2020. Discussion Paper 18905. CEPR. URL: https://cepr.org/system/files/publication-files/DP18905.pdf.

Carlsson, B., Stankiewicz, R., 1991. On the nature, function and composition of technological systems. Evolutionary Economics 1, 93–118.

Dehler-Holland, J., Okoh, M., Keles, D., 2022. Assessing technology legitimacy with topic models and sentiment analysis – the case of wind power in Germany. Technological Forecasting and Social Change 175, 121354. doi:10.1016/j.techfore.2021.121354.

DellaVigna, S., Ferrar, E.L., 2015. Economic and Social Impacts of the Media. Working Paper 21360. NBER .

für Wirtschaft und Energie, B., 2019. Freiräume für Innovationen. Das Handbuch für Reallabore.

Gentzkow, M., Shapiro, J.M., 2010. What drives media slant? evidence from U.S. daily newspapers. Econometrica 78, 35–71. URL: https://www.jstor.org/stable/25621396.

Grimmer, J., Stewart, B.M., 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Analysis 21, 267–297. doi:10.1093/pan/mps028.

Grubb, M., Drummond, P., Poncia, A., McDowall, W., Popp, D., Samadi, S., Penasco, C., Gillingham, K., Smulders, S., Glachant, M., et al., 2021. Induced innovation in energy technologies and systems: A review of evidence and potential implications for CO2 mitigation. Environmental Research Letters 16, 43007.

IFR, 2023. Executive Summary World robotics 2023 — Service Robots. Technical Report. International Federation of Robotics (. URL: https://ifr.org/img/worldrobotics/Executive_Summary_WR_Service_Robots_2023.pdf.

Kang, I., Yang, J., Lee, W., Seo, E.Y., Lee, D.H., 2023. Delineating development trends of nanotechnology in the semiconductor industry: Focusing on the relationship between science and technology by employing structural topic model. Technology in Society 74, 102326. doi:10.1016/j.techsoc.2023.102326.

Kelly, B., Papanikolaou, D., Seru, A., Taddy, M., 2021. Measuring technological innovation over the long run. American Economic Review: Insights 3, 303–320. doi:10.1257/aeri.20190499.

Köper, M., Schulte im Walde, S., 2016. Automatically generated affective norms of abstractness, arousal, imageability and valence for 350 000 German lemmas, in: Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S. (Eds.), Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), European Language Resources Association (ELRA), Portorož, Slovenia. pp. 2595–2598. URL: https://aclanthology.org/L16-1413.

Krebs, F., Peller-Konrad, F., Younes, A., Reister, F., Bärmann, L., Vetter, P., Weinberger, N., Loewe, M., Ott, I., Nierling, L., Asfour, T., 2023. Making AI Tangible for Children through Humanoid Robots, in: International Conference on Child-Robot Interaction.

Lehotský, L., Černoch, F., Osička, J., Ocelík, P., 2019. When climate change is missing: Media discourse on coal mining in the Czech Republic. Energy Policy 129, 774–786. doi:10.1016/j.enpol.2019.02.065.

Loewe, M., Quittkat, C., Knodt, M., Ott, I., 2024. The Impact of the Russian War against Ukraine on the German Hydrogen Discourse. Sustainability 16, 773. doi:10.3390/su16020773.

Nierling, L., Weinberger, N., Vetter, P., Maia, M.J., Asfour, T., Krebs, F., Peller-Konrad, F., Reister, F., Younes, A., Bärmann, L., Loewe, M., 2023. Holding the Tension between Technological Pathways and Societal Needs : A critical Reflection on a Technology Oriented Real-World Lab, in: 21st Annual STS Conference : Critical Issues in Science, Technology and Society Studies.

Ott, I., forthcoming. TA in innovation systems, in: Grunwald, A. (Ed.), International Handbook of Technology Assessment. Edward Elgar. chapter 31.

Roberts, M.E., Stewart, B.M., Airoldi, E.M., 2016a. A model of text for experimentation in the social sciences. Journal of the American Statistical Association 111, 988–1003. doi:10.1080/01621459.2016.1141684.

Roberts, M.E., Stewart, B.M., Tingley, D., 2016b. Navigating the local modes of big data: The case of topic models, in: Alvarez, R.M. (Ed.), Computational Social Science. Cambridge University Press, pp. 51–97. doi:10.1017/cbo9781316257340.004.

Roberts, M.E., Stewart, B.M., Tingley, D., 2019. stm: An R package for structural topic models. Journal of Statistical Software 91, 1–40.

Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S., Albertson, B., Rand, D., 2014. Structural topic models for open-ended survey responses. American Journal of Political Science 58, 1064–1082.

Rudenko, I., Norman, U., Maure, R., Rudenko, A., Weinberger, N., Krebs, F., Peller-Konrad, F., Asfour, T., Bruno, B., 2024. Drawings for Insight on Preschoolers' Perception of Robots, in: ACM/IEEE International Conference on Human Robot Interaction (HRI).

Savin, I., Ott, I., Konop, C., 2022. Tracing the evolution of service robotics: Insights from a topic modeling approach. Technological Forecasting and Social Change 174, 121280. doi:10.1016/j.techfore.2021.121280.

Schäpke, N., Stelzer, F., Caniglia, G., Bermann, M., Wanner, M., Singer-Bodowski, M., Loorbach, D., Ollson, P., Baedecker, C., Lang, D.L., 2018. Jointly experimenting for transformation? shaping real-world laboratories by comparing them. GAIA 27/S1, 85–96.

Sohns, A., 2023. Differential exposure to drinking water contaminants in north carolina: Evidence from structural topic modeling and water quality data. Journal of Environmental Management 336, 117600. doi:10.1016/j.jenvman.2023.117600.

Zhang, W., Cao, G., Ji, Y., Gu, L., Wang, S., 2022. Analysis of electric vehicle technology development based on patent big data: a topic analysis of structured topic model (STM), in: Zhao, F. (Ed.), 5th International Conference on Computer Information Science and Application Technology (CISAT 2022), International Society for Optics and Photonics. SPIE. p. 124512H. URL: https://doi.org/10.1117/12.2656586, doi:10.1117/12.2656586.

Table 4
Topic Scores for Four Psycholinguistic Attributes ($\in [0, 10]$), Topical Prevalence Contrast ($\in [-0.07, 0.13]$), and Eigenvector Centrality ($\in [0, 1]$). Own calculation based on the full text corpus and disregarding timeline.

Topic Label		Valence	Arousal	Abstr.ness	Imag.blty	Contrast	Evt
Government Funding for Research on AI and Climate Change	1	5.206	3.955	4.065	4.205	0.013	0.633
Fair Pay Innovation Lab	2	5.044	3.977	3.827	3.917	0.020	0.621
Machine Learning and AI Development Methods	3	5.203	4.025	3.615	3.779	-0.068	0.539
Artificial Humans in Movies and Literature	4	5.236	4.284	3.564	4.080	-0.051	0.449
Digitization of Business Processes	5	5.036	3.942	3.616	3.626	-0.024	0.617
Chatbots and ChatGP T	6	4.851	3.769	3.699	3.627	-0.031	0.623
AI and Robotics Cutting-Edge Research Funding in Bavaria	7	4.885	3.943	3.957	4.137	0.011	0.704
Robotics and AI in Construction, Agriculture and Policing	8	5.239	3.872	4.195	4.428	0.019	0.000
Real-World Labs for the Energy Transition	9	5.209	3.860	3.983	4.088	0.089	0.000
Tightening of Investment Controls in the High-Tech Sector (USA, EU, Germany)	10	4.857	4.069	3.922	3.983	-0.029	0.436
Digitalization in Urban Development	11	5.171	3.846	4.217	4.511	0.037	0.059
Regular Airspace Pilot Projects in Germany	12	5.059	3.813	3.971	4.085	0.047	0.014
Covid Mass Vaccination as a Field Experiment	13	4.753	4.069	3.421	3.669	0.018	0.476
Google	14	4.557	3.636	3.974	3.862	-0.032	0.786
Future Lab in Aachen - the Whole City as a Future Lab	15	5.429	4.000	4.225	4.504	0.080	0.000
Tech Start-Ups	16	5.049	3.791	3.967	3.983	0.008	0.436
Autonomous Cars	17	4.691	3.901	4.261	4.335	-0.032	0.313
Real-World Labs for Sustainable Mobility in Baden-Württemberg	18	5.394	3.840	4.053	4.227	0.055	0.178
Robots and AI in Private Households	19	5.162	3.797	4.396	4.408	-0.027	0.187
Service Robots and Assistance Systems	20	5.162	4.024	3.877	4.100	-0.014	0.279
Relationships between Humans, Machines and AI in Art Exhibitions	21	5.463	4.258	4.155	4.668	-0.027	0.345
Robots in Production	22	4.983	3.757	4.115	4.112	-0.035	0.527
Digitization, Robots and AI in Schools	23	5.514	4.273	4.205	4.480	0.000	0.650
Digital Transformation Leaders	24	5.052	3.859	3.775	3.864	0.016	0.388
Softbank (Japanese Tech Investor)	25	4.820	3.847	3.903	3.941	-0.035	0.344
Digitalization and Automation Consequences for the Workplace	26	4.930	4.209	3.668	3.913	-0.046	0.646
Political Support for AI and in Germany	27	5.279	4.152	3.684	3.928	-0.021	0.847
Real-World Labs for Ecofriendly Mobility in Aachen	28	4.942	3.851	3.924	4.154	0.126	0.057
Robots in Space and Robot Development	29	5.095	4.010	4.144	4.288	-0.022	0.128
Philosophical Considerations on the Digital Transformation	30	5.558	4.561	3.458	4.067	-0.010	1.000
Lethal Autonomous Weapon Systems	31	4.726	4.545	3.553	3.845	-0.034	0.641
Fintech and Legal Tech	32	5.007	3.745	3.848	3.813	-0.001	0.573