

Shkolnykova, Mariia

Article — Published Version

Who shapes plant biotechnology in Germany? Joint analysis of the evolution of co-authors' and co-inventors' networks

Review of Evolutionary Political Economy

Provided in Cooperation with:

Springer Nature

Suggested Citation: Shkolnykova, Mariia (2020) : Who shapes plant biotechnology in Germany? Joint analysis of the evolution of co-authors' and co-inventors' networks, Review of Evolutionary Political Economy, ISSN 2662-6144, Springer International Publishing, Cham, Vol. 2, Iss. 1, pp. 27-54,
<https://doi.org/10.1007/s43253-020-00027-y>

This Version is available at:

<https://hdl.handle.net/10419/288786>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Who shapes plant biotechnology in Germany? Joint analysis of the evolution of co-authors' and co-inventors' networks

Mariia Shkolnykova¹ 

Accepted: 21 October 2020 / Published online: 14 November 2020

© The Author(s) 2020

Abstract

The interdependence of science and technology has been of high interest for researchers from different fields for several decades now. As they represent different means of knowledge output protection, patents and research articles generally have different reasons for creation and different audiences. However, some of the inventors may be interested in making an impact on the scientific community and vice versa. This interaction between technology space (patents) and science space (articles) is especially important for high-technology fields, where both research institutions and enterprises play important roles in the innovative environment. This paper investigates the interaction between science and technology in the case of German plant biotechnology. With the help of network analysis tools, the evolution and co-evolution of co-inventors' and co-authors' networks for the period 1995–2015 is explored. Finally, the topics of the patents and papers from the overlap were analyzed with the help of text mining tools in order to identify the differences of topics between science/technology and their overlap. As a result, sizable differences in nature and advancement are observed between the two network types. Although the overlap between these spheres of innovative activities increased over time, the role author-inventors played in the science or technology space varied.

Keywords Technological space · Science space · Network analysis · Biotechnology · Patents · Bibliometrics

JEL classification O31 · O34 · D85

✉ Mariia Shkolnykova
mariia.shkolnykova@uni-bremen.de

¹ Faculty of Business Studies and Economics, University of Bremen, Bremen, Germany

1 Introduction

Since its appearance in the 1990s, German biotechnology has received much attention from different parties: policymakers, university researchers, prominent corporations, and SMEs. Funding initiatives (BioRegio, BioProfile, BioChance¹) have boosted the entries of firms in the industry as well as the patenting activity of existing and new actors (Adelberger 2000; Kaiser and Prange 2004; Müller 2002). Also, research institutions are among the main actors of the field (Casper et al. 1999; Müller 2002). Apart from that, many dedicated biotechnology firms in Germany have been founded as academic spin-offs (Simmie 2002). Therefore, besides patents, publications are also an important form of protecting intellectual property rights within the biotechnology landscape (e.g., Katila 2000).

Although patents and publications both indicate the flow of innovations, they represent two different “worlds,” with patents reflecting the commercialization of inventions and publications reflecting the conceptual and intellectual side (Noyons et al., 1998). Therefore, several papers have used a combination of data of both patents and research articles in order to analyze innovative activities in the field. Different fields appear to have varying degrees of science-technology overlap: ranging from a rather large overlap for industries such as biotechnology and pharmacy to a rather small overlap for industries such as tissue engineering (Murray 2002). Such conclusions, however, are often made after having only considered the non-patent references of patent literature (Meyer 2002; Verbeek et al. 2002; Guan and He 2007; Breschi and Catalini 2010) or after having performed a comparative analysis of the publication and patent data pointing out the most important keywords (De Looze 1994). Moreover, previous papers have not looked at the evolution of both networks to see how interactions have developed over time and to identify pioneers in the field, as well as to determine whether science and technology tend to converge or diverge as the industry evolves. The advancements of the role of author-inventors across both science and technology spaces are far from being understood. Furthermore, the analysis of thematic fields, which are important for both industry and academia, is also limited.

This paper analyzes the case of plant biotechnology in Germany. The choice of this sub-field and region can be explained by its controversial nature, caused by the debate on genetically modified organisms. Because of this debate, plant biotechnology is considered underdeveloped in Germany (McCormick and Kautto 2013), despite its high potential in addressing sustainability issues. The time scope for the study is 1995 to 2015, which reflects the time from the initiation of biotechnology in Germany until the year when full number of patents could be identified. Thus, this exploratory paper deals with the following research questions: (1) How did networks of co-authors and co-inventors develop over time? (2) How many actors are active both in science and technology for the case of German plant biotechnology? (3) What role do author-inventors play within both networks and how did this role change over time? (4) How do the topics of the author-inventors differ from just-authors and just-inventors?

¹ These initiatives were initiated by Federal Ministry of Education and Research (BMBF) with an aim to promote biotechnology SMEs (BioChance), create biotechnology clusters (BioRegio), or support regions with a specific profile (BioProfile). More information about these initiatives can be found on the website of BMBF: <https://www.bmbf.de/>.

To address these questions, the paper proceeds as follows. First, the data from two independent datasets on co-authors and co-inventors are separately analyzed over sequential periods of time. Then the data is merged in order to find the overlap, showing actors who are active in both the scientific and the technological sphere as well as the ones who only engage in one type of activity. For the overlap, the centrality measures of actors are calculated in order to identify what influence they have on the whole network. Furthermore, text mining tools allow for the identification of topics that were important for both spaces, independently as well as simultaneously.

This paper provides several important contributions in regard to methodology as well as to the existing literature. Methodologically, the paper extends the ways of identification of patents and articles belonging to a specific sub-field, based on a keyword search procedure. Furthermore, it supports the discourse on the interdependencies between science and technology by providing a long-term observation and presenting the trends of science-technology interaction over time. In addition, it contributes to the development of matching procedures when identifying the population of author-inventors.

To extend the existing literature, this paper is the first to provide information about the science-technology overlap for the case of the highly disputable field of plant biotechnology. Thus, it complements other works, dealing with the science-technology interplay in biotechnology in general (Breschi and Catalini 2010), as well as in such related fields as pharmacy (McMillan et al. 2000) and tissue engineering (Murray 2002). Furthermore, based on previous studies, the paper provides information on how the role of author-inventors in German plant biotechnology has changed over time. Thus, instead of providing a static picture, the paper looks at the dynamics of the science-technology overlap as well as the changes in author-inventor roles across both networks separately. This allows to make assumptions about what lies behind this change—the influence of legislation, the trajectory of field development, or policy incentives.

The paper is structured as follows: Section 2 provides the theoretical background and an overview of existing literature on the topic of science-technology interaction, and additionally outlines the peculiarity of biotechnology in this regard. Section 3 includes the presentation of both technology and science datasets as well as the joint dataset and provides descriptive statistics. Section 4 introduces the research methodology, including network and text mining analysis techniques. Moreover, the procedure for identifying matches between co-inventors' and co-authors' networks is presented. Section 5 presents the results of the network analysis over sequential time periods on the level of the entire network, on the node level, as well as the author-inventors' network. This is followed by the text analysis of the overlap. The paper concludes with the discussion of the existing limitations and implications for different parties.

2 State of the art

2.1 Theoretical background

Initially, science and technology were perceived as independent spaces or “universes” (Coward and Franklin 1989). Academia was seen as the provider of knowledge and

“abstract” research, whereas industry was responsible for commercialization and technical solutions. However, over time, the interdependencies and overlaps between the two spaces were growing and the line between them became harder to draw (Murray 2002). This tendency was supported by the concept of the triple helix: the relationship between university, industry, and governance, which outlines the importance of academia for industry development and vice versa (e.g., Meyer et al. 2003; Bhattacharya and Meyer 2003). Thus, several authors saw the need to investigate how scientists contribute to patenting, which is normally perceived as an industry prerogative, and how enterprise engineers and top managers contribute to paper publications in scientific journals (Balconi et al. 2004; Breschi and Catalini 2010; Meyer 2002; Murray 2002). There are certain incentive mechanisms, which lie behind academia being interested in patenting and vice versa. Universities and research institutions are among the main actors not only in the research space, but also in the technological space via engaging and pushing productive collaborations as well as providing grounds for spin-offs and subsidiaries (Roesler and Broekel 2017). This is especially true for high-tech industries, which are dependent on high-quality research on different levels (Pammolli and Rossi 2005). Another incentive for the emergence of a broad range of author-inventors lies in the disappearance of the difference between basic and applied research. This leads to university-industry knowledge transaction: on the one hand, firms are also investing in basic research via academic publishing, and, on the other hand, universities more and more often extend their intellectual property management strategies to include patents (Pammolli and Rossi 2005). This ambivalent behavior allows both firms and research institutions to stay alert to the new developments along both research types. One more trend, which leads to author-inventors’ appearance, lies in interdisciplinarity, which is prevailing across many fields, especially new disciplines (e.g., bioinformatics, nanotechnology, ICT). This requires existence and co-existence of various actors as well as different types of copyright, including research articles, patents, and trademarks (Pammolli and Rossi 2005; Eisenberg 1999). This variety is necessary in order to fulfill one of the main functions of intellectual property: the transactional function, i.e., to enable knowledge exchange and knowledge access by as many actors as possible (Pammolli and Rossi 2005).

Another type of incentive—the promoting of patenting activities by research institutions—is given by the legislative initiatives of many countries. In the USA, for example, the change started with the Bayh–Dole Act, which removed restrictions for university patenting and thus resulted in the establishment of transfer offices in universities, allowing them to commercialize and license their intellectual property (Tahmooresnejad and Beaudry 2019). Similar legislation was applied in Canada at the end of the 1990s (Tahmooresnejad and Beaudry 2019). A number of studies discuss the positive impact of such institutional settings on university patenting (e.g., Aldridge and Audretsch 2011; Audretsch 2014; Link and Siegel 2005). However, this impact is only notably present in the USA (Tahmooresnejad and Beaudry 2019). In Europe, the situation was slightly different. In many countries, the inventors of the patents (not the universities) were seen as their owners.² In Germany, this legislation was in place until 2002 (Schoen et al. 2014). This legislation also led to many patents being filed by enterprises, especially start-ups (Schoen et al. 2014). These findings suggest that when analyzing science-technology interactions, the inventor level has to be taken into account.

² This was also called “professor’s privilege.”

In the literature, the identification of author-inventors and their contribution to technology was predominantly based on non-patent references of patents (e.g., Guan and He 2007; Breschi and Catalini 2010, Zhang et al. 2019). Other papers tried building two independent datasets for patents and academic publications, and combining them with the help of different matching techniques (e.g., Klitkou et al. 2007; Boyack and Klavans 2008; Gautam et al. 2014). Both paths started with identifying relevant patent- or patent and publication populations with the help of keyword search or the use of particular technological categories and research topics.

The ways in which the obtained datasets were approached methodologically also differ. Many authors explored the interaction of science and technology by applying social network analysis (e.g., Breschi and Catalini 2010; Zhang et al. 2019) and clustering techniques (e.g., Chang et al. 2010; Zhu et al. 2019). In that respect, the calculation of centrality measures and outlining them for author-inventors was the primary interest (e.g., Balconi et al. 2004; Zhang et al. 2019). Additionally, some researchers used the networks to build maps of keywords (Madani and Weber 2016; Wang et al. 2018). In some cases, the received data served as variables for regression analysis (e.g., Beaudry and Kananian 2012) or even to forecast emerging technologies (Daim et al. 2006).

In this paper, the existing literature is used and extended to get a comprehensive picture of the science-technology interaction. First, the co-evolution of two networks (co-inventors and co-authors) over sequential time periods is investigated. Also, certain techniques, presented in former studies, are applied to follow the network positions of actors from the overlap (author-inventors) and how they developed over time. Furthermore, this paper expands the usage of text mining techniques for the analysis of science-technology interaction by identifying the main topics from the overlap of the two networks, as well as the ones that are only relevant for either science or technology.

2.2 Special case of biotechnology

Most papers that deal with the interplay between science and technology focus on a specific industry or research/technological field. Authors like Meyer (2006), Rueda et al. (2007), and Tahmooresnejad and Beaudry (2019) focused on the field of nanotechnology. De Looze (1994), Dalpé (2002), and Glänzel and Zhou (2011) turned their attention to biotechnology. Noyons et al. (1998) concentrated on the IT sector, Klitkou et al. (2007) on fuel cells, and Murray (2002) on tissue engineering research. Furthermore, Breschi and Catalini (2010) covered several sectors with their analysis: lasers, semiconductors, and biotechnology. Thus, the research is usually concentrated on high-tech industries. One reason for this is that in such industries, many university-industry collaborations and innovations can be expected. Outside the IT sector, patents can also be seen as the common way of protecting inventions in these industries and thus can serve as a measure for creating technology space (e.g., Baum et al. 2000; Katila 2000).

The above-stated papers show diverse results with respect to the grade of overlap between science and technology. Whereas for biotechnology (Breschi and Catalini 2010), fuel cells (Klitkou et al. 2007), and pharmacy (McMillan et al. 2000), the intersection is usually reported to be quite big, for other industries like tissue engineering (Murray 2002), it is relatively small. Breschi and Catalini (2010) also report that

matched actors (inventors who are also authors) appear to have a more central position in the network than non-matched ones (just-inventors or just-authors).

This paper focuses on the case of German plant biotechnology.³ The reasons for this are manifold. Firstly, plant biotechnology is an important part of the field due to its high potential of addressing sustainability issues such as global hunger by increasing plant yield (Brooks 2005). However, it is still underdeveloped in Germany due to legislation issues and its controversial nature (McCormick and Kautto 2013); therefore, it would be interesting to learn more about its development processes. Secondly, the importance of scientists for the development of biotechnology has been explicitly stated by previous research (e.g., Zucker and Darby 1996). Therefore, it would be interesting to find out whether the impact of the research community will be strengthened or weakened over time. Thirdly, Kudic and Shkolnykova (2020) show the technological trajectory of biotechnology goes in the direction of cell tissue research and wound care, which, as mentioned before, have rather low science-technology overlap. This contrasts with large science-technology overlap, generally expected for biotechnology (Breschi and Catalini 2010). Therefore, the grade as well as the dynamics of the number of author-inventors and their network positions requires more research.

The time scope of the paper is 1995–2015. Thus, all stages of biotechnology development in Germany are covered: starting with its origins in the mid-1990s, when there were several government funding initiatives (e.g., BioRegio, BioFuture,⁴ BioProfile, BioChance), through to the end of this funding period and the relative downturn in the mid-2000s. In the late 2000s and early 2010s, there was a transition in the field with respect to both present business models and the technological scope. Apart from that, 2015 is the year for which a complete number of patents could be obtained.⁵

3 Dataset creation and descriptive statistics

3.1 Technological space

The creation of the technological space starts with identifying German plant biotechnology patents and creating networks of co-inventors for their population. In order to create an initial patent population among biotechnology classes⁶ according to the International Patent Classification (IPC), plant-related classes were chosen based on the initial screening and number of appearances of the word “plant” within patent titles and abstracts.⁷

³ According to the National Institute of Food and Agriculture: “Plant biotechnology is a set of techniques used to adapt plants for specific needs or opportunities” (as stated on <https://nifa.usda.gov/plant-biotechnology> Last Accessed on 30 July 2020).

⁴ BioFuture initiative was initiated by BMBF as a platform for young researchers’ support (Müller 2002).

⁵ With the help of the EPO Worldwide Patent Statistical Database (PATSTAT), Version 2017b.

⁶ IPC Classes: A01 H1/00, A01 H4/00, A61 K38/00, A61 K39/00, A61 K48/00, C02 F3/34, C07 G, C07 K, C12 M, C12 N, C12 P, C12 Q, C12 R, C12 S, G01 N27/327, G01 N33/ (53 *, 54 *, 55 *, 57 *, 68, 74, 76, 78, 88, 92) as well as later added C40B 40/00 -50/18, C40B 70/00 -80/00, C40B 10/00.

⁷ Resulting classes: A01H 1*, A01H 4*, A61K 38/56, C07K 4/10, C07K 14/415, C07K 16/16, C12M 3*, C12N 5*, C12N 9/32, C12N 15/05, C12N 15/29, C12N 15/82, C12Q 1/6895.

After that, patents from the specified technology classes with the priority filing year between 1995 and 2015 and having at least one inventor or applicant from Germany were identified. Furthermore, in order to substantially reduce the number of strategic patents, only those patent families that have filings in German and European patent offices or Patent Cooperation Treaty (PCT) were included in the sample.⁸ As a result, 2075 patent families were identified with 10,648 inventor entries.⁹ Furthermore, these entries were cleaned from duplicates: entries with the same inventor names were deleted, if they had the same address or/and same co-inventors. As the result, only 2437 inventor entries were left. It shows that plant biotechnology in Germany is created by a relatively small group of inventors. The resulting entries correspond to 8881 pairs of co-inventors, which were further used for the creation of the technology space.

3.2 Science space

To obtain the co-authors' network, the *Web of Science* database,¹⁰ which is the most common bibliometric source for such kind of research (Tan et al. 2014), has been used. As *Web of Science* usually presents papers from peer-reviewed journals, some time passes between the first submission and publication. Thus, following Dornbusch et al. (2013), the population of publications was taken with 1-year lag, i.e., 1996–2016.¹¹ This helps in better identifying the point in time, when the new scientific idea was created.

In comparison to IPC, *Web of Science* categories are rather broad; therefore, a deeper usage of keyword search methods was needed in order to identify relevant (plant biotech) papers. For that, the literature suggests using an initial list of words, which is later developed and specified using citation networks (Mogoutov and Kahane 2007). Bentley (2008) used four influential articles in the field of social and natural sciences and looked at the patterns of keywords of citing articles in order to follow the evolution of academic vocabulary in specific fields. Therefore, to build a vector of keywords reflecting plant biotechnology, the following procedure was used. In the first step, one of the most prominent papers on plant biotechnology was taken as a starting point for the search for an initial list of keywords, which are relevant for plant biotechnology. This paper was chosen to be *Gill. S. S. and Tuteja. N. (2010): "Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants."* This paper has 5953 citations^{12,13} and is one of the most influential papers in the category "Plant Science" according to the *Web of Science* database. Also, it is located in the 99th percentile according to citations on Scopus. In the next step, citations of

⁸ Strategic patents in this case mean that a firm (or a person) applies for a patent in a specific country in order to get a monopoly power on a particular invention without a specific intention to produce there. Because of different policies of intellectual property rights protection, many strategic patents can be applied for in the USA or China, however, never come to the world (PCT) or European level. Apart from that, a focus on German, European, and world level allows to reflect the German technology space.

⁹ Based on *psn_id* variable along *tis209_appl_ipc* table, which reflects PATSTAT standardized name.

¹⁰ Database, which provides a wide range of scientific journal articles, books, and proceedings from different research fields. It can be accessed under <http://isiknowledge.com/wos>.

¹¹ For simplicity, 1995–2015 is stated overall in the analysis part.

¹² According to Google Scholar, retrieval date 04.02.2020.

¹³ Theoretically, every well-cited and plant biotechnology relevant paper could be taken at this step. The focus here lies on the most occurring keywords in the citing papers, which appeared to be similar for several often cited papers in the field.

this paper are analyzed and the list of keywords, stated in the focal paper (Gill S.S. & Tuteja N. 2010) as well as in the citing papers, is created. This step is necessary, as according to Bentley (2008), the selection of keywords and their diversification is higher for natural sciences (in the case of Bentley (2008)—physics) than for social sciences. It means that it will not be enough to look only at the keywords of the focal article in order to build a list of keywords that covers articles related to plant biotechnology. Thus, the most popular keywords (more than 70 occurrences with elimination of general keywords, e.g., “article,” “stress,” “concentration”) were then chosen from focal and citing papers’ keywords. The resulting 129 words were then used to create a population of plant biotechnology articles.¹⁴

Finally, the query on *Web of Science* was performed to generate the scientific space. Resulting papers should include the identified words as well as match several additional filters: (1) publication year between 1996 and 2016, (2) subject areas include “Biotechnology and Applied Microbiology”, and (3) at least one author has German affiliation. Here, 21,474 papers with 102,875 author entries were identified. These entries were further checked for duplicates based on the same full name and affiliation. Apart from that, as the scope of the paper is limited to German science and technology overlap, the authors without German affiliation were deleted from the population. As a result, 21,781 authors were left in the network, which correspond to 310,484 pairs of co-authors over the whole period.

3.3 Joint space

As could already be seen above, science space of plant biotechnology appeared to be far more extensive than technology space. Figure 1 shows that the number of papers exceeds the number of patents at least four times each year. Furthermore, whereas within science space an almost constant upward trend was observed, within technology space, the number of patents decreases over time. As a result, in 2015, the number of papers exceeds the number of patents about 20 times.

With regard to the number of inventors and authors per year, different tendencies can be observed. Here as well, the number of authors exceeds the number of inventors substantially each year. However, the number of inventors appears to be relatively stable over time, whereas the number of authors is fluctuating. It generally has an upwards trend but a considerable decrease in 2008–2009, followed by an increase until 2016, when another decrease occurred. A possible explanation for the downturns could be a general radical shift in the field, also observed for the patent data (Kudic and Shkolnykova 2020), as well as the end of many funding initiatives, or even the world economic crisis, which could have had influence on research financing. Also, at that time, many smaller actors experienced changes in the business model towards service orientation (Kahl 2015).

As the aim of the paper is to examine how the interaction between science and technology developed over time, the matched author-inventors had to be identified along subsequent periods of time. Therefore, three 5-year periods and one 6-year period were set: 1995–1999; 2000–2004; 2005–2009; and 2010–2015.¹⁵ These periods correspond approximately to the stages of the industry development: the 1st period relates

¹⁴ These words are presented in Appendix A.

¹⁵ For science space, 1-year lags were taken into account.

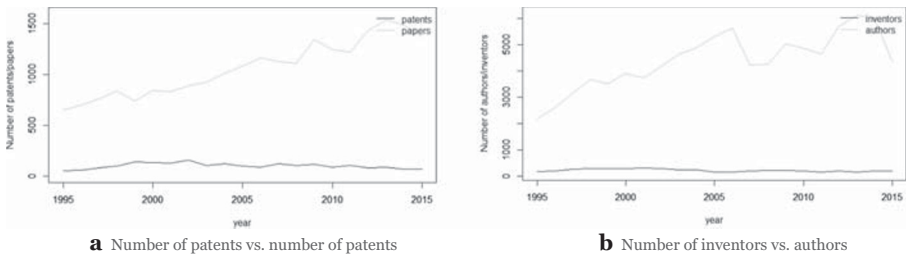


Fig. 1 Descriptive statistics, science, and technology space

to the emergence of biotechnology in Germany, the 2nd corresponds to its growth, the 3rd to the decline in the number of patents, and the end of the period for most of the governmental funding programs, and the 4th period reflects the time of re-profiling of the industry (Kudic and Shkolnykova 2020).

After both datasets were created, matching procedure was started in order to identify the overlap between the science and technology space. It is based on existing literature (Coward and Franklin 1989; Boyack and Klavans 2008; Schoen et al. 2014; Dornbusch and Neuhäusler 2015, Dornbusch et al. 2013) and applies the following procedure:

- *1st step*—Matching was performed on the basis of identical name and surname. Standard typos and errors, occurring in PATSTAT (e.g., writing of German “Umlauts” with Unicode characters) could be automatically solved; however, there were also some typing errors in inventors’ names that needed to be manually checked and corrected.
- *2nd step*—Using the dataset, resulting from the 1st step, an additional revision was performed based on the same affiliation of authors and inventors. Thus, possible false-positive matches could be identified. Whereas in the case of research articles affiliation is usually explicitly stated, in the case of technology space, the applicant of a corresponding patent serves as a proxy for the inventor’s affiliation. It means, however, that not for every inventor the affiliation could be identified, especially for the case of European inventors in the 1990s, because of the “professor’s privilege.” Therefore, an additional filtering step was needed.
- *3rd step*—Co-authors and co-inventors of the matched candidates from the 1st step were scanned. In case the same co-authors and co-inventors could be identified, the match was perceived to be correct. However, following the Boyack and Klavans (2008) definition of rare names,¹⁶ this verification was done only for very common names (e.g., Schwarz, Müller) in the sample.

Contrary to the method of Dornbusch et al. (2013), the locations of authors and inventors were not taken into account. The reason for this is that here not only academic inventions and research articles are analyzed. Therefore, it cannot be expected that all inventors live in the same region with the enterprises or institutions, stated in the patent documents as applicants, especially for the case of MNEs’ patents.

¹⁶ In this case, name is identified only with one institution is associated with author/inventor, or only a few organizations are associated with one inventor with high strong dominance of one organization across cases.

As the result of the procedure, 652 matches could be identified for the period 1995–2015, 116 for the period 1995–1999, 156 matches for 2000–2004, 191 for the period 2005–2009, and 228 for 2010–2015, with several matches identical for several periods. This number corresponds to around 1% of the total number of authors and 17–18 to 27–29% of inventors. Whereas the number of matched author-inventors¹⁷ remains relatively stable over time, the number of inventor-authors experienced a significant increase during the third observation period. As this period was also marked by the declining number of inventors in general, a first conclusion may be that many of these inventor-authors stayed in the sample, whereas just-inventors dropped out.

As can be seen from the affiliation diagram of matched nodes (Fig. 2), no clear dominant institution can be seen across author-inventors. Predictably, most of them come from universities and research institutions. BASF AG is the most popular organization among author-inventors, being the affiliation for ten nodes.

4 Methodology

4.1 Social network analysis procedure

Methodologically, the paper is based on the social network analysis techniques (e.g., Wasserman and Faust, 1994; Borgatti et al. 2018), which have their origins in graph theory. According to this methodology, each of the three datasets (co-authors, co-inventors, and matched dataset) was organized in the network form with nodes¹⁸ being individual actors (co-authors or co-inventors) and edges¹⁹ being present whenever nodes share the same patent (for co-inventors' network) or the same paper (for co-authors' network). Following Borgatti et al. (2018), this paper includes several levels of analysis: the network level, the node level, and the level of matched networks.

4.1.1 Analysis on the network level

In order to get the impression of the overall network, several standard measures, well accepted in the literature (e.g., Fritsch and Kudic 2019, Borgatti et al. 2018), were applied. This is done in order to observe the dynamics of the network as an entire system, reflected in the science and technology space (Fritsch and Kudic 2019). All measures were calculated separately for the same subsequent periods: 1995–1999; 2000–2004, 2005–2009, and 2010–2015.

To start with, the basic indicators, reflecting network structure and size, are calculated. These include the total number of nodes and edges engaged in the network in each period, as well as the average number of edges (average degree) among all nodes in a particular period. Thus, it can be observed whether the network expands or shrinks over time. These measures also provide the first insights into its connectivity. Furthermore, the share of isolates (nodes, which do not have any edges) was estimated. This

¹⁷ In this case, the term author-inventors means that the actors appeared in science space before technology space, whereas inventor-authors mean that the actors appeared in technology space before science space. For convenience, throughout the paper, the term “author-inventor” is used to describe both types of actors.

¹⁸ Also known as vertices

¹⁹ Also known as link, dyads, or ties

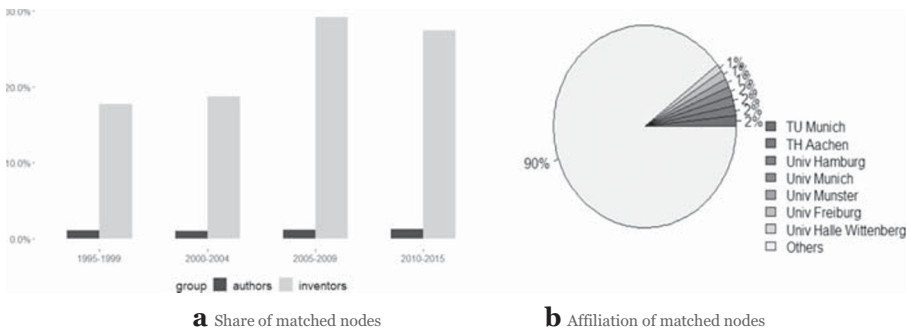


Fig. 2 Descriptive statistics, matched nodes

indicator shows the proportion of the actors, which do not have any ties. For the here-presented science and technology space, it means the number of authors or inventors, who do not have any collaborator with German affiliation.

Additionally, several other measures were chosen to describe network connectivity. These help to gain deeper insights into the size of the groups of researchers working together and how well-linked they are among each other. To identify connected groups, the notion of (weak) component is used. Weak component is a sub-graph, where each node can reach every other node (Borgatti et al. 2018). Using this definition, for each network, the size of the three biggest components is presented for each observation period. This helps to define the biggest connected group, which either works on a patent or a paper together, or several such groups, connected by at least one tie. Relating the measure of the biggest component size to the overall network size makes it possible to show the ratio of authors or inventors connected to the main group, and thus able to reach each other. Whenever every network actor can reach the other actors (either through neighbors or through own ties), the graph can be described as a “small world network” (Borgatti et al. 2018; Wasserman and Faust, 1994).

However, when talking about network structure, the interest lies often not only in the largest components. Therefore, the component ratio is analyzed at this stage. It is calculated as the ratio between the number of components and the number of nodes. Mathematically, the indicator can be calculated as follows (Perry et al. 2018):

$$CR = \frac{K-1}{N-1} \quad (1)$$

where K is the number of components and N is the number of nodes in the network.

The component ratio reflects the connectedness of the network, with 1 meaning that none of the authors or inventors has a collaborator affiliated in Germany (number of components is equal to the number of nodes) and 0 meaning that all authors or inventors in the network are connected (K is equal to 1) (Borgatti et al. 2018).

As the focus of the paper lies on the changes of the network structures and of the positions of specific nodes (author-inventors), the dynamics of actors within each network is observed. The measures used for this purpose are the shares of the new and remaining nodes. New nodes were defined as the ones not present in the network one period before. The remaining nodes are the ones that also existed in the network

one period before. Also, the share of reoccurring nodes was calculated for periods 3 and 4. These are the nodes, which, although not present in period $t-1$, were in the network in period $t-2$ or $t-3$. The same measures were examined on the edge level (following e.g. Broekel and Bednarz 2018).

4.1.2 Analysis on the node level

The analysis of the properties of individual nodes was performed with a set of standard centrality measures (e.g., Borgatti et al. 2018; Zhang et al. 2019; Breschi and Catalini 2010; Wanzenböck et al. 2013). These allow an inference about the position of a particular node in the network as well as about the importance each node has in the network structure (Borgatti et al. 2018).

First, degree centrality was identified as the simple number of edges that a node has.²⁰ Thus, here the importance of the node can be reflected just through the number of connections with co-authors or co-inventors. In this case, authors or inventors with the highest number of ties may be seen as the most visible in the community (Borgatti et al. 2018). Second, betweenness centrality was calculated as the number of shortest paths that go through the node. This indicator helps to identify the so-called gatekeepers of the network: without them, clusters will break (Borgatti et al. 2018). In the case of authors or inventors, actors with a high level of betweenness centrality connect two otherwise unconnected working groups, e.g., former employees of one department shifting to another one. These actors may also bring along knowledge, acquired while producing an article or invention, and share it with their connected actors.

Mathematically, betweenness centrality can be described as:

$$C_B(i) = \sum_{j \neq i, k \in n} \frac{g_{jk}(i)}{g_{jk}} \quad (2)$$

where $g_{jk}(i)$ is the number of the shortest paths between j and k and g_{jk} is the number of the shortest paths between j and k , which go through i .²¹

Another important measure is eigenvector centrality, which reflects the importance of a particular node within the network. Like degree centrality, this measure reflects the number of ties the node has. However, connected nodes are weighted based on their centralities, which means that the node can only be as central as his partners are (Borgatti et al. 2018). For the case of co-inventors' or co-authors' networks, this measure helps to identify actors who are linked to the most popular authors or inventors.

Mathematically, in order to calculate this measure, a matrix is used with cell $a_{ij} = 1$ if there is a connection between i and j , and $a_{ij} = 0$ if there is no connection between these nodes:

$$C_e(i) = \lambda \sum_{j=1}^n a_{ij} x_j \quad (3)$$

where a is the eigenvector of the adjacency matrix A with the eigenvalue λ .

²⁰ As all networks of these paper are undirected—the direction of edges is not of interest—no distinction between in-degree and out-degree measures was made.

²¹ Another important measure, closeness centrality, was not calculated in this paper, as there can occur problems when interpreting such measure for disconnected networks.

4.1.3 Analysis on the matched network level

For the case of the matched network, all general network measures are calculated. Furthermore, the centrality characteristics of the matched nodes for co-inventors' and co-authors' networks are estimated. It includes the estimation of their share among top inventors and top authors. Thus, it can be identified which properties the actors in the overlap have—are they usually well-connected and central actors or do author-inventors usually find themselves on the network periphery? Following the example of Breschi and Catalini (2010), the analysis is supported by investigating, whether differences between matched and non-matched samples can be found with the help of the Mann–Whitney–Wilcoxon test. This non-parametric test, based on the ranks of the observations, allows identifying whether matched nodes have the same positions according to different centrality measures as just-authors or just-inventors. The analysis is performed in RStudio²² with the help of codes, presented in e.g. Luke (2015).

4.2 Text mining applications

After performing the network analysis, text mining techniques are used in order to identify the main topics along the matched networks as well as for co-authors' and co-inventors' networks. This allows to show the topics that have importance only for science or only for technology as well as the ones that are relevant for both fields. Apart from that, as the keywords for subsequent periods may differ, the analysis may also help to show how the topics have developed over time. The inputs for the analysis were patent and paper titles.²³ They provide the key idea of the scientific or technological output. Only papers and patents in English language are taken into account. Also, the dataset was restricted in order to delete stop words, plurals, and numbers.

In the last step, according to Silge and Robinson's (2017) codes for RStudio, the most co-occurring keywords were selected both for non-matched and matched actors and visualized based on the frequencies of the co-occurrences. Thus, it was possible to generate a picture of the field and identify separate clusters of connected keywords. By comparing the most occurring keywords in matched and non-matched nodes' networks, it can be seen how authors, inventors, and author-inventors differ regarding their research fields.

5 Results

5.1 Science and technology space analysis on the network level

Thus, the overall network measures were calculated for both co-inventors' and co-authors' networks over all periods. First, the results for the co-inventors' networks are presented.

²² The open-source software for R, which can be accessed under <https://rstudio.com/>.

²³ Other alternative types of input that could be used for text-mining analysis in the case of papers are *Author keywords* or *KeyWords Plus keywords*, generated by *Web of Science*. For the *KeyWord Plus* case, keywords are generated from titles of the references, cited by the focal article. However, for consistency reasons, in this paper, titles were used for both papers and patents.

Figure 3 shows co-inventors' networks for four sequential periods. Considering the overall network structure, apart from the last period, one large component followed by a large number of smaller ones can be observed. Also, the share of isolates is relatively high, at one point even surpassing 20%. This shows a significant number of single-person inventions. Generally, the degree of centralization of the network is low, reflecting a modest level of centralization and a decentralized structure.

With regard to the changes of the network structure over time, it can be seen that, in the end of the 1990s, one big cluster existed in the network followed by many smaller cliques of inventors. Not unusual was also the situation of just two inventors working and patenting together. In the second and third period, several other clusters of inventors appeared, which grew with time, whereas the first component diminished in size. In the last period, starting from 2010, several clusters of co-inventors can be observed, without any cluster dominating in size. Thus, the general trend seems to go from big teams towards smaller ones, which goes along the line of the trend from big towards small firms in plant biotechnology. The following analysis of network measures will show more insights from these networks.

Table 1 shows the values obtained on the overall network level for the co-inventors' network. As could already be seen from the figure, the network first tends towards consolidation, with the component ratio getting smaller and their size increasing, and then towards separation, with many small components appearing. This is also reflected in the diminishing average length of the path and decreased network diameter. In general, it can be said that with regard to inventions, the industry experienced structural changes in the mid-2000s, which were followed by a re-orientation and a new wave of inventors coming to the field in the 2010s. This statement can also be supported by the

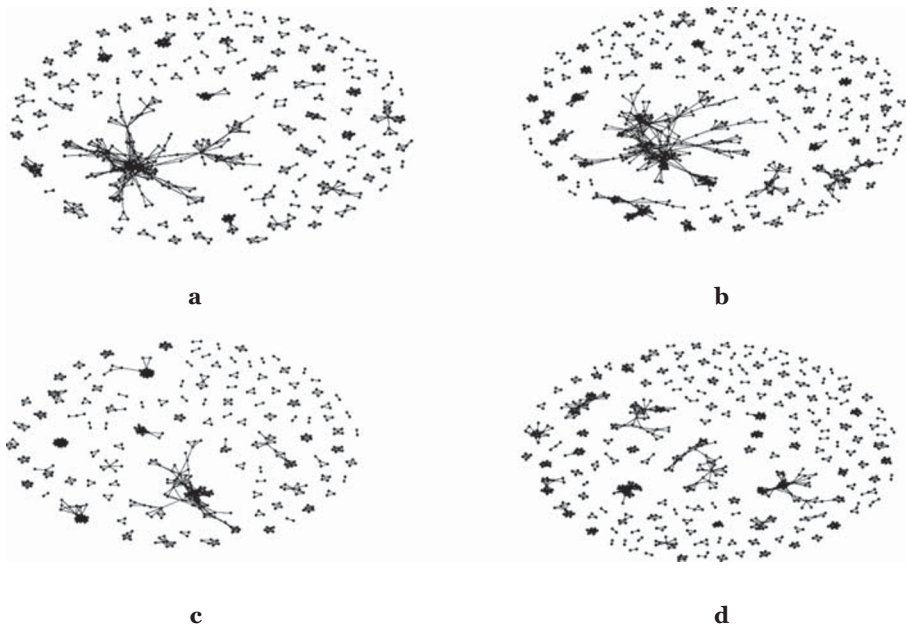


Fig. 3 Co-inventors' networks over time. **a** Co-inventors' network, 1995–1999. **b** Co-inventors' network, 2000–2004. **c** Co-inventors' network, 2005–2009. **d** Co-inventors' network, 2010–2015

number of nodes, remaining in the network. It can be seen that the proportion of nodes, remaining or reoccurring in the network, changes over time. Whereas in the third observation period this number is growing, and between 2005 and 2009 each 4th node has been seen in the network before, in the fourth period, this number falls to less than 20%. It shows that the 2010s were marked by new inventors, who either followed the changing technological trend within biotechnology and changed their field of patenting or just created their first patent.

Interestingly, the number of remaining or reoccurring edges is approximately 2.5–3 times less than that of the nodes. It means that even the nodes, remaining in the network, mostly have changed their co-inventors. This situation may especially often occur within corporations with large development teams or in big institutions, which mostly work on projects. Notably, in the last period, the number of reoccurring edges is 2% higher than the number of remaining edges. It means that some nodes revived their co-inventorship from the beginning of the 2000s.

Next, the same measures for the network of co-authors were created. Figure 4 presents the development of co-authors' networks over time.

For the co-authors' network, one constantly growing large component, followed by many really small ones (only around 1% of the largest component regarding the number of nodes), can be observed. Furthermore, contrary to the co-inventors' network, the share of the isolates is quite low here, even though slightly growing over time, which means that the number of entries with a single German author is quite small. The low degree centralization parameter, as for the case of co-inventors, shows a low degree of network: even though several central nodes are present, almost all actors share a relatively similar number of ties.

Table 1 Co-inventors' overall network measures

	1995–1999	2000–2004	2005–2009	2010–2015
Number of nodes	653	833	656	872
Number of edges	1498	2100	1435	1677
Share of isolates	0.081	0.102	0.216	0.157
Degree centralization	0.068	0.062	0.053	0.043
Average length of the path	3.911	3.843	2.874	2.268
Network diameter	10	10	9	8
Number of components	105	124	100	151
Component ratio	0.172	0.148	0.151	0.172
Size of the biggest component	155	179	75	48
<i>Size of the 2nd biggest component</i>	19	38	25	38
<i>Size of the 3rd biggest component</i>	14	37	23	29
New nodes	–	0.813	0.780	0.859
Remaining nodes	–	0.187	0.220	0.141
Reoccurring nodes	–	0.187	0.252	0.197
New edges	–	0.937	0.916	0.939
Remaining edges	–	0.063	0.084	0.061
Reoccurring edges	–	0.063	0.085	0.081

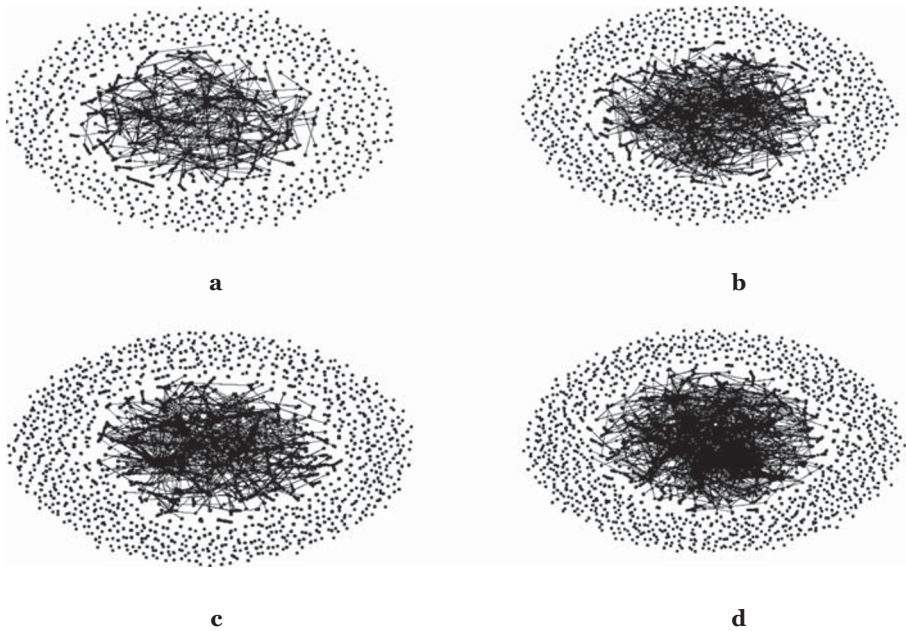


Fig. 4 Co-authors' networks over time. **a** Co-authors' network, 1995–1999. **b** Co-authors' network, 2000–2004. **c** Co-authors' network, 2005–2009. **d** Co-authors' network, 2010–2015

In comparison to co-inventors' network dynamics, in the case of co-authors, an increase in the size of the main component and a general increase in the number of nodes can be observed. No other large components, comparable to the first one, appear in the network. Thus, additional exploration is important here, which can be provided by introducing the overall network measures (Table 2). Here, the tendencies are different from those of the co-inventors' networks. Namely, the main component, which composes about $2/3$ of the overall number of nodes, grows over time, with the number of components and isolates increasing only slightly; thus, the network rather converges. It means that a large proportion of nodes, coming to the network over time, are added to the main component, or the existing nodes start to collaborate with the main components' participants. The lower diameter and average path length as well as significantly increased number of edges show that new ties are appearing between previously unconnected actors, which is mostly probable for the main path.

Additionally, this tendency can be seen in the large number of new edges. At least 94.4% of the network edges have not occurred in the network before. This means that only a small number of authors maintain their co-authorship ties. It especially contrasts with the number of nodes, remaining or reoccurring in the network, which is growing over time and is reaching in 2010–2015 more than 38%.

Thus, the first conclusion, which can be made here, is that co-authors' and co-inventors' networks have different development paths. Whereas the co-inventors' network has a trend towards divergence and dissimulation of the main component, the network of co-authors experiences a constant growth of the main component, which includes most of the network, whereas the 2nd largest component is more than 100 times smaller. This means that the co-authors' network is mainly concerned with one

Table 2 Co-authors' overall network measures

	1995–1999	2000–2004	2005–2009	2010–2015
Number of nodes	10,838	15,183	16,457	18,777
Number of edges	53,351	65,578	68,769	73,959
Share of isolates	0.006	0.004	0.023	0.039
Degree centralization	0.017	0.012	0.011	0.009
Average length of the path	8.187	7.920	7.666	6.952
Network diameter	24	21	21	20
Number of components	613	797	964	1044
Component ratio	0.057	0.050	0.059	0.056
Size of the biggest component	7028	10,208	10,328	12,862
<i>Size of the 2nd biggest component</i>	76	98	68	69
<i>Size of the 3rd biggest component</i>	55	47	67	51
New nodes	–	0.823	0.835	0.783
Remaining nodes	–	0.177	0.165	0.217
Reoccurring nodes	–	0.177	0.181	0.383
New edges	–	0.985	0.999	0.946
Remaining edges	–	0.015	0.001	0.054
Reoccurring edges	–	0.015	0.006	0.056

topic (with probably many subtopics) developing over time, whereas in the co-inventors' network, several almost equally important components are being developed.

5.2 Science and technology space analysis on the node level

After the analysis on the level of the entire network, individual characteristics of nodes are examined. The summary of node centrality measures for the case of co-inventors' networks is presented in Table 3. Here, isolates were not included in the analysis. Corresponding to the results of the overall network analysis, the average number of ties, which a node possesses, first increases and then decreases, which is due mostly to the diminishing size of the main component. The same may be said about other centrality measures—the network is becoming more even over time, with less nodes being crucially different from the others.

The same measures were estimated for the case of co-authors' networks. The results of this analysis are presented in Table 4. Here again, the trend goes in the opposite direction from that of co-inventors'. The average betweenness centrality increases over time; however, this happens mostly because of several outliers. Apart from that, average degree centrality falls, showing that even central nodes have less direct connections over time.

Thus, the analysis on the node level additionally supports the finding of differences between co-inventors' and co-authors' networks. Whereas the co-inventors' network is becoming more even, in the network of co-authors, central authors are getting more important over time.

Table 3 Co-inventors' summary of centrality measures

	Min.	Median	Mean	Max.
1995–1999				
Degree centrality	1.000	3.000	4.960	46.000
Betweenness centrality	0.000	0.000	63.620	4506.08
Eigenvector centrality	0.000	0.000	0.025	1.000
2000–2004				
Degree centrality	1.000	4.000	5.395	52.000
Betweenness centrality	0.000	0.000	71.050	3458.290
Eigenvector centrality	0.00	0.000	0.024	1.000
2005–2009				
Degree centrality	1.000	4.000	5.591	33.000
Betweenness centrality	0.000	0.000	15.690	896.270
Eigenvector centrality	0.000	0.000	0.037	1.000
2010–2015				
Degree centrality	1.000	3.000	4.454	37.000
Betweenness centrality	0.000	0.000	7.198	462.000
Eigenvector centrality	0.000	0.000	0.016	1.000

5.3 Matched network analysis

In the next step, the author-inventors' network was analyzed in regard to the presence of matched nodes in the main component as well as in the top actors according to different centrality measures. Following Breschi and Catalini (2010), to show the relative importance of author-inventors in the overall network, the top 25% of the nodes, as defined by different centrality characteristics, were examined. Then, the ration of the share of author-inventors in the top 25% of the nodes to the share of all author-inventor nodes in the population was calculated. Thus, for example, in the period from 1995 to 1999 out of 653 inventors in general,²⁴ 116 were matched nodes, which correspond to a share of 0.178. Of these matched nodes, 30 were among the top 25%, which correspond to a share of 0.184. The representation of author-inventors in the top 25% nodes for that case is 1.034, which means that they are slightly overrepresented in the top nodes. Appendix B presents the results of the same analysis for the top 10% and top 5% of nodes.

In contrast to the findings of Breschi and Catalini (2010), Table 5 does not show univocal results with regard to the representation of matched nodes under top-central authors and inventors. Only for the co-authors' network, matched nodes are overrepresented in almost all centrality categories. Here, the representation of the matched nodes across top-central actors is getting bigger over time. Only for the case of eigenvector centrality in the first two periods no overrepresentation can be stated. It shows that author-inventors, starting on average on almost similar positions like only authors, are getting more and more influence in the network over time.

²⁴ Here, isolates were also taken into account, as they can be matched as well.

Table 4 Co-authors' summary of centrality measures

	Min.	Median	Mean	Max.
1995–1999				
Degree centrality	1.000	6.000	9.908	194.000
Betweenness centrality	0.000	0.000	16,494	3,813,498
Eigenvector centrality	0.000	0.000	0.014	1.000
2000–2004				
Degree centrality	1.000	6.000	8.677	193.000
Betweenness centrality	0.000	0.000	23,860	5,370,201
Eigenvector centrality	0.00	0.000	0.0038	1.000
2005–2009				
Degree centrality	1.000	6.000	8.559	190.000
Betweenness centrality	0.000	0.000	22,134	5,011,370
Eigenvector centrality	0.000	0.000	0.004	1.000
2010–2015				
Degree centrality	1.000	6.000	8.194	176.000
Betweenness centrality	0.000	0.000	27,275	10,523,253
Eigenvector centrality	0.000	0.000	0.004	1.000

The situation in the co-inventors' network is different. Here, no specific overrepresentation of the nodes can be seen across almost all periods and measures, with the only exception of the betweenness centrality in 1995–1999. Apart from that, the representation of inventor-authors across all top categories diminishes over time. Thus, one conclusion here may be that the majority of matched inventors could not keep their central positions and could not adapt well to the transformation of technology space. Another possible reason for underrepresentation of inventor-authors may be that after the field transformation started, many researchers decided to focus on scientific outputs.

Because of these vague results, the question arises whether the centrality measures of the matched actors differ significantly from those of non-matched. Table 6 provides the results of the Mann–Whitney–Wilcoxon test that was performed in order to see whether there are differences in the ranks between the matched and non-matched

Table 5 Representation of author-inventors, top 25% of nodes

	1995–1999	2000–2004	2005–2009	2010–2015
Matched nodes, top degree centrality, co-inventors	1.034	1.003	0.984	0.773
Matched nodes, top degree centrality, co-authors	1.107	1.238	1.600	1.829
Matched nodes, top betweenness centrality, co-inventors	1.206	1.080	0.984	0.949
Matched nodes, top betweenness centrality, co-authors	1.711	1.554	1.843	2.255
Matched nodes, top eigenvector centrality, co-inventors	0.896	1.028	0.754	1.142
Matched nodes, top eigenvector centrality, co-authors	0.671	0.869	1.539	1.314

group. The test was performed following Breschi and Catalini (2010) and chosen because of the high skewness of distribution.

The results show that there are almost no differences between matched and non-matched samples for co-inventors. Thus, it cannot be said that inventor-authors possess more central positions in the network than their counterparts. However, the opposite can neither be stated. Thus, inventor-authors are rather evenly distributed across the network. The only significant result obtained for co-inventors relates to the measure of eigenvector centrality in the two later periods. This means that inventor-authors, although not being the most influential nodes across the network themselves, appear to have highly connected co-inventors. This may relate to the situation where university researchers are invited or commissioned by a firm to realize some industrial project.

A different result is found for matched author-inventors. After the first two periods, matched nodes possess significantly higher centrality. This result additionally shows differences in the development path of science and technology space: whereas within science, authors who have technical know-how usually occupy influential positions in the network, inventors, active in publications, are spread evenly across the network.

Appendix C shows the location of matched nodes across co-inventors' and co-authors' networks over time. For the first three periods in the co-inventors' networks, not many matched nodes found themselves in the main component. These are rather small cliques of three or four inventors that could be also found as authors. However, as around 1/5–1/3 of the nodes can be found among the top actors according to centrality characteristics, it can be stated that the ones that find themselves in the main component

Table 6 Results of Mann–Whitney–Wilcoxon test

	Mean, matched	Mean, inventors	Mean, authors	<i>p</i> value
1995–1999				
Degree centrality	5.441, 9.250	4.863	9.915	0.954 0.9962
Betweenness centrality	64.578, 47,619.300	63.426	16,166.730	0.294 0.000***
Eigenvector centrality	0.007 0.009	0.029	0.014	0.498 0.009***
2000–2004				
Degree centrality	5.503, 8.299	5.368	8.681	0.987 0.333
Betweenness centrality	61.240, 47,122.97	73.455	23, 620.64	0.497 0.000***
Eigenvector centrality	0.007, 0.000	0.027	0.004	0.236 0.023**
2005–2009				
Degree centrality	5.626, 11.769	5.578	8.525	0.324 0.001***
Betweenness centrality	13.645, 77,373.57	16.456	21, 550.71	0.743 0.000***
Eigenvector centrality	0.036, 0.000	0.038	0.005	0.078* 0.005***
2010–2015				
Degree centrality	4.288, 13.106	4.508	8.136	0.763 0.000***
Betweenness centrality	11.712, 10,7953.63	5.738	26, 334.67	0.447 0.000***
Eigenvector centrality	0.007, 0.009	0.019	0.004, 0.004	0.049** 0.000***

p value column: first value—only inventors, second value—only authors

***Rejection of the null hypothesis on 0.01 significance level, **Rejection of null hypothesis at 0.05 significance level, *Rejection of the null hypothesis on the 0.1 significance level

play an important role. The situation in the fourth period, however, is different. Although author-inventors there do not dominate over the shrinking main component (taking 1/3 of its nodes), they are well represented in one of the other big components in the network (fourth biggest component with 27 nodes in total, of which 16 are inventor-authors). This means that the same project team, which files a patent can be found on the authors' list. In the co-authors' network, the situation is relatively stable. The share of nodes in the main component is getting bigger along with the component itself.

Finally, the stability of the nodes in the overlap is checked. For that, it is specifically looked at, which fraction remains or reoccurs in both co-authors' and co-inventors' networks (Table 7). Generally, the remaining matched nodes are overrepresented across networks. This number, however, is unstable and different for actors' and inventors' networks. Whereas the percentage of new matched nodes in the co-inventors' network for the last period corresponds approximately to the overall network indicator, for the case of co-authors, the number of remaining and reoccurring matched nodes is highly overrepresented.

5.4 Text mining tools' application

In order to dig deeper into the nature of differences between the two types of networks, the content of patents and articles was analyzed. This can show if the main ideas studied by matched author-inventors correspond to the ones dominant for the field in general or if they are completely different. To do so, titles presenting the main idea of the invention or scientific product were collected and scrutinized. The most co-occurring bigrams were visualized in networks with nodes presenting individual words, with the width of the edge reflecting the frequency of co-occurrences of specific word pairs. Figure 5 presents results of the keyword analysis for the co-inventors' network for the overall period 1995–2015. The results for the sequential periods are presented in Appendix D.

From the first look at the main and non-main component keyword networks, they may seem similar as both have the topic of transgenic plants in the core. However, the subtopics are different along the main and non-main components. Whereas in the main component the topic of increased yield as well as plant tolerance and resistance is dominant, the non-main component deals with topics of cell and tissue research. When looking at the matched sample, it seems that they are thematically more connected to the topics of the main component. Thus, these are mostly the actors who try to follow the dominant stream of

Table 7 Stability across matched nodes

	1995–1999	2000–2004	2005–2009	2010–2015
New nodes, co-inventors	–	0.571	0.749	0.794
Remaining nodes, co-inventors	–	0.429	0.251	0.206
Reoccurring nodes, co-inventors	–	0.429	0.288	0.293
New edges, co-authors	–	0.500	0.686	0.520
Remaining edges, co-authors	–	0.500	0.314	0.480
Reoccurring edges, co-authors	–	0.500	0.325	0.511

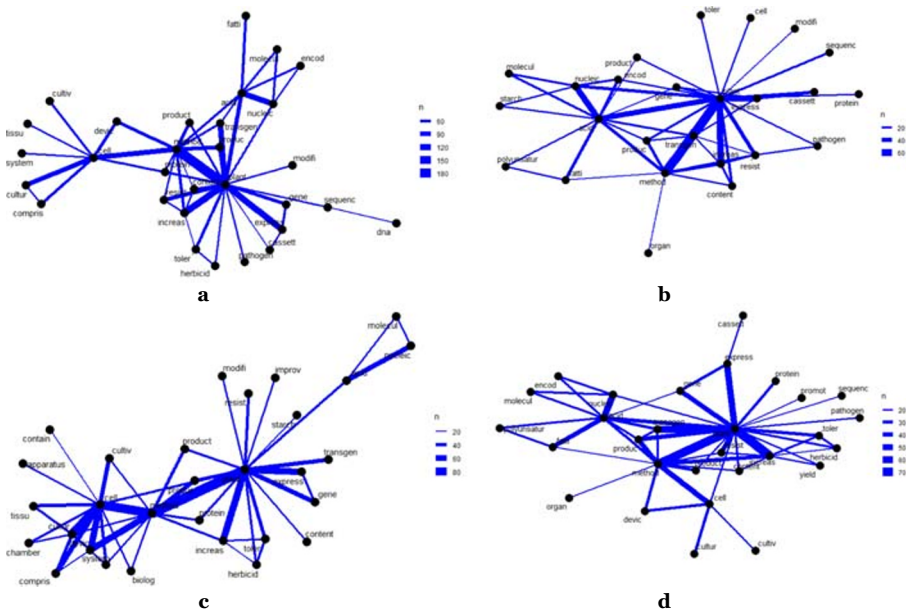


Fig. 5 Keyword network, co-inventors. **a** Complete network. **b** Main component, non-matched. **c** Non-main component, non-matched. **d** All matched

thoughts across the field, even though they are not always connected to the nodes of the main component or do not have central positions in it. The topics of the non-main component are also present; however, they are not dealt with as actively as with transgenic plants' research.

Figure 6 presents the same exercise now done for the network of co-authors over the period 1995–2015, whereas Appendix E presents the changes in keyword networks occurring over time. Here, along the first three networks, the focus lies on medical biotechnology as well as on applications of specific bacteria for plant modification and DNA/RNA recombination (a topic, which is also outlined by the industry). Particularly, whereas along the main components the topics are mostly of general nature and concerning some bacteria, commonly used in biotechnology research, non-main components focus on specific terms. With regard to the matched component, two streams can be identified here: (1) nodes that are publishing in the same field in which they are patenting (e.g., in the topic of transgenic plants, dealing thus with the same practical issue of increased yield and resistance), which does not have popularity across either main or non-main components; (2) nodes, dealing with these topics indirectly through the terms like “*Escherichia coli*,” or “*Corynebacterium glutamicum*,” that have usage across many other fields, including different medical implications. These nodes, thus, perform more classic university research. Additional explorations show that the topics of the second stream come also from the nodes situated in the main component. The ones dealing with transgenic plants directly are located mostly on the network's periphery, in small components, and thus, dispatched from the core of science space.

but the number of inventor-authors was continuously growing. Probably, not only the increase in the overall number of inventors causes this rising number of inventor-authors but also the increased interdisciplinarity of the field (Pammolli and Rossi 2005) and the investigation of medicine-related topics (Kudic and Shkolnykova 2020), which allows a higher number of actors to engage in the scientific discussion. Another explanation might be the intensified industry-university collaborations, which lead to researchers of applied science and industry representatives to be included as co-authors (Roesler and Broekel 2017).

Referring to the third research question regarding the role of author-inventors in both networks during different periods, their overrepresentation in the most central nodes, shown by previous research (Breschi and Catalini 2010), could be supported only for the case of the co-authors' network. Furthermore, the value of the centrality measure of author-inventors in the co-authors' network even increases over time. For the network of co-inventors, almost no differences were identified between matched and non-matched actors. This means that the author-inventors could take a central place in the research landscape but not in the inventive one.

Lastly, regarding the fourth research question considering the differences of topics between matched and non-matched nodes, the disparities were also identified when looking at the topics, which the author-inventors are dealing with. Two streams of actors could be identified. On the one hand, they are those who deal with the topics of transgenic plants and who are active in the co-inventors' network and evenly spread among components there. Also, in the technological space, they are not highly represented across top-central nodes, and in the science space, these actors also do not have the most central positions. On the other hand, those actors could be identified, who deal with more general topics, which could be also applied in medicine, and who are not located in the main component of co-inventors but provide the basis of the science space. Nevertheless, it is important to mention the transformation occurring within the plant biotechnology field (also outlined by Kudic and Shkolnykova 2020). The nodes outside the main component are getting more and more connected. The continuation of this study with actualized data (e.g., for 1995–2020) will help to unveil whether the shift towards blending plant and medicine-related biotechnology continues.

The paper has therefore several important implications. First, for policy-makers it outlines the importance of funding for research institutions, which may create impact for the research and academic field. Specifically, in the scientific network matched actors are of special importance. Second, for actors themselves, it shows the importance of collaboration and diversification. Those actors who perform research and apply for patents in different subfields of plant biotechnology appear to be located more central among authors and are also regarded in inventors' networks. Third, it represents the transformation patterns of the field itself, which is of high importance for all stakeholders. In particular, the medical orientation may be of special interest in solving societal issues, such as providing health and global well-being. The interplay of basic and applied research may be of particular importance in this case.

Nevertheless, this paper still has some limitations and potential for further research. First, although trying to account for the case of "professor's privilege," in some cases, the shift of the rights on the patent from the university professor to the university may diminish the individual motivation to patent. Apart from that, further insights into the importance of specific topics and their development over time could be traced by following the changes in ownership of particular patents.

Moreover, several limitations and further research possibilities could be identified regarding the employed data. First, there could have been several false-negative entries, as some inventors or authors could have changed their surname for some reason and therefore were not stated as author-inventors. It could especially influence the number of nodes, which remain in the matched network over time. However, it is supposed that these situations do not happen often enough to change the results drastically.

Furthermore, this paper is focused on the analysis of cooperation networks of authors and inventors from Germany. Thus, results regarding disparities of both types of networks can be applied only to the German case. However, Germany is an influential country with regard to innovation activities (for both patent and research activities), which is based on a broad institutional support of both academia (BioFuture) and industry (BioChance, BioRegio) (Müller 2002). Thus, the experiences and trends of German science-technology interaction may serve as an example for other countries. Further research could investigate the cross-country interactions of science and academia. It may be of special interest because of the various institutional and legislative environments different countries are subjects to.

Another important point, which can extend existing research, may lie in creating two-layer networks (following e.g. Lucena-Piquero and Vicente 2019), thus separating between connections, occurring within an institution and the ones, occurring between institutions. This may help correcting several network measures for the possible biases for the case of authors or inventors, belonging to well-connected project teams. For this study, such a network could not be created, because of the difficulties of attaching inventors to a particular organization. However, further research and development of matching algorithms may help to overcome this limitation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s43253-020-00027-y>.

Acknowledgments The author is indebted to the participants of the 31st European Association for Evolutionary Political Economy (EAEPE) Conference and 5th EAEPE Research Area [X] “Networks” Workshop for helpful comments on the earlier versions of this paper and the German Academic Exchange Service (DAAD) for providing travel grant to visit EAEPE Conference. The author also wants to thank anonymous reviewers for providing valuable feedback and comments.

Code availability Network analysis and text mining analysis were performed with the help of RStudio, open source software. Codes were modified based on the information from Luke (2015) and Silge and Robinson (2017) and are available on request. Initial explorations of the dataset were performed in MS Excel.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author upon request.

Compliance with ethical standards

Conflict of interest The author declares that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adelberger KE (2000) Semi-sovereign leadership? The state's role in German biotechnology and venture capital growth. *Ger Polit* 9(1):103–122. <https://doi.org/10.1080/0964400008404582>
- Aldridge TT, Audretsch D (2011) The Bayh–Dole Act and scientist entrepreneurship. *Res Policy* 40(8):1058–1067. <https://doi.org/10.1016/j.respol.2011.04.006>
- Audretsch DB (2014) From the entrepreneurial university to the university for the entrepreneurial society. *J Technol Transf* 39(3):313–321. <https://doi.org/10.1007/s10961-012-9288-1>
- Balconi M, Breschi S, Lissoni F (2004) Networks of inventors and the role of academia: an exploration of Italian patent data. *Res Policy* 33(1):127–145. [https://doi.org/10.1016/S0048-7333\(03\)00108-2](https://doi.org/10.1016/S0048-7333(03)00108-2)
- Baum JA, Calabrese T, Silverman BS (2000) Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. *Strateg Manag J* 21(3):267–294. [https://doi.org/10.1002/\(SICI\)1097-0266\(200003\)21:3<267::AID-SMJ89>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<267::AID-SMJ89>3.0.CO;2-8)
- Beaudry, C., & Kananian, R. (2012) Impact of university–industry contracts resulting in patents on the quality of patenting in biotechnology. DRUID 2012 Conference, Copenhagen, Denmark. https://conference.druid.dk/acc_papers/frugyobi5utbpm0c30slkvaxraj.pdf.
- Bentley RA (2008) Random drift versus selection in academic vocabulary: an evolutionary analysis of published keywords. *PLoS One* 3(8):e3057. <https://doi.org/10.1371/journal.pone.0003057>
- Bhattacharya S, Meyer M (2003) Large firms and the science/technology interface—patents, patent citations, and scientific output in thin films. *Scientometrics* 58(2):265–279. <https://doi.org/10.1023/A:1026284510104>
- Borgatti SP, Everett MG, Johnson JC (2018) *Analyzing social networks*, 2nd edn. SAGE Publications Ltd., New York
- Boyack KW, Klavans R (2008) Measuring science–technology interaction using rare inventor–author names. *J Informetr* 2(3):173–182. <https://doi.org/10.1016/j.joi.2008.03.001>
- Breschi S, Catalini C (2010) Tracing the links between science and technology: an exploratory analysis of scientists' and inventors' networks. *Res Policy* 39(1):14–26. <https://doi.org/10.1016/j.respol.2009.11.004>
- Broekel T, Bednarz M (2018) Disentangling link formation and dissolution in spatial networks: an application of a two-mode STERGM to a project-based R&D network in the German biotechnology industry. *Netw Spat Econ* 18(3):677–704. <https://doi.org/10.1007/s11067-018-9430-1>
- Brooks S (2005) Biotechnology and the politics of truth: from the green revolution to an evergreen revolution. *Sociol Rural* 45(4):360–379. <https://doi.org/10.1111/j.1467-9523.2005.00310.x>
- Casper S, Lehrer M, Soskice D (1999) Can high-technology industries prosper in Germany? Institutional frameworks and the evolution of the German software and biotechnology industries. *Ind Innov* 6(1):5–24. <https://doi.org/10.1080/13662719900000002>
- Chang PL, Wu CC, Leu HJ (2010) Using patent analyses to monitor the technological trends in an emerging field of technology: a case of carbon nanotube field emission display. *Scientometrics* 82(1):5–19. <https://doi.org/10.1007/s11192-009-0033-y>
- Coward HR, Franklin JJ (1989) Identifying the science–technology interface: matching patent data to a bibliometric model. *Sci Technol Hum Values* 14(1):50–77. <https://doi.org/10.1177/016224398901400106>
- Daim TU, Rueda G, Martin H, Gerdts P (2006) Forecasting emerging technologies: use of bibliometrics and patent analysis. *Technol Forecast Soc Chang* 73(8):981–1012. <https://doi.org/10.1016/j.techfore.2006.04.004>
- Dalpé R (2002) Bibliometric analysis of biotechnology. *Scientometrics* 55(2):189–213. <https://doi.org/10.1023/A:1019663607103>
- De Looze MA (1994) The application of scientometric tools to the analysis of a sector in plant biotechnologies: nitrogen fixation. *Scientometrics* 30(1):23–34. <https://doi.org/10.1007/BF02017210>

- Dornbusch, F., & Neuhäusler, P. (2015) Academic patents in Germany (No. 6-2015) Studien zum deutschen Innovationsystem. <http://hdl.handle.net/10419/156616>.
- Dornbusch F, Schmoch U, Schulze N, Bethke N (2013) Identification of university-based patents: a new large-scale approach. *Res Eval* 22(1):52–63. <https://doi.org/10.1093/reseval/rvs033>
- Eisenberg R (1999) Bargaining over the transfer of proprietary research tools: is this market failing or emerging. In: Cooper Dreyfuss R (ed) Expanding the boundaries of intellectual property. Oxford University Press, Oxford
- Fritsch, M., & Kudic, M. (2019) Micro fluidity and macro stability in inventor networks (No. 2019-004). Jena Economic Research Papers. <http://hdl.handle.net/10419/204608>.
- Gautam P, Kodama K, Enomoto K (2014) Joint bibliometric analysis of patents and scholarly publications from cross-disciplinary projects: implications for development of evaluative metrics. *J Contemp East Asia* 13(1):19–37. <http://hdl.handle.net/2115/56810>
- Glänzel W, Zhou P (2011) Publication activity, citation impact and bi-directional links between publications and patents in biotechnology. *Scientometrics* 86(2):505–525. <https://doi.org/10.1007/s11192-010-0269-6>
- Guan J, He Y (2007) Patent-bibliometric analysis on the Chinese science—technology linkages. *Scientometrics* 72(3):403–425. <https://doi.org/10.1007/s11192-007-1741-1>
- Kahl, J. (2015) The determinants of firm growth in high-technology industries. An empirical analysis of spatial, relational and institutional factors in the German biotechnology industry. Schriftenreihe innovative betriebswirtschaftliche Forschung und Praxis, Band 443, Hamburg
- Kaiser R, Prange H (2004) The reconfiguration of national innovation systems—the example of German biotechnology. *Res Policy* 33(3):395–408. <https://doi.org/10.1016/j.respol.2003.09.001>
- Katila R (2000) Using patent data to measure innovation performance. *Int J Bus Perform Manag* 2(1/2/3):180–193. <https://EconPapers.repec.org/RePEc:ids:ijbpm:v:2:y:2000:i:1/2/3:p:180-193>
- Klitkou A, Nygaard S, Meyer M (2007) Tracking techno-science networks: a case study of fuel cells and related hydrogen technology R&D in Norway. *Scientometrics* 70(2):491–518. <https://doi.org/10.1007/s11192-007-0213-6>
- Kudic, M.; Shkolnykova, M. (2020) From biotech to bioeconomy: new empirical evidence on the technological transition to plant-based bioeconomy based on patent data. Bremen Papers on Economics & Innovation. No. 2002. <https://elib.suub.uni-bremen.de/edocs/00108535-1.pdf>. Accessed 27 February 2020
- Link AN, Siegel DS (2005) University-based technology initiatives: quantitative and qualitative evidence. *Res Policy* 34(3):253–257. <https://doi.org/10.1016/j.respol.2005.01.005>
- Lucena-Piquero D, Vicente J (2019) The visible hand of cluster policy makers: an analysis of Aerospace Valley (2006-2015) using a place-based network methodology. *Res Policy* 48(3):830–842. <https://doi.org/10.1016/j.respol.2019.01.001>
- Luke DA (2015) A user's guide to network analysis in R. Springer, Cham
- Madani F, Weber C (2016) The evolution of patent mining: applying bibliometrics analysis and keyword network analysis. *World Patent Inf* 46:32–48. <https://doi.org/10.1016/j.wpi.2016.05.008>
- McCormick K, Kautto N (2013) The bioeconomy in Europe: an overview. *Sustainability* 5(6):2589–2608. <https://doi.org/10.3390/su5062589>
- McMillan GS, Narin F, Deeds DL (2000) An analysis of the critical role of public science in innovation: the case of biotechnology. *Res Policy* 29(1):1–8. [https://doi.org/10.1016/S0048-7333\(99\)00030-X](https://doi.org/10.1016/S0048-7333(99)00030-X)
- Meyer M (2002) Tracing knowledge flows in innovation systems. *Scientometrics* 54(2):193–212. <https://doi.org/10.1023/A:1016057727209>
- Meyer M (2006) Are patenting scientists the better scholars?: an exploratory comparison of inventor-authors with their non-inventing peers in nano-science and technology. *Res Policy* 35(10):1646–1662. <https://doi.org/10.1016/j.respol.2006.09.013>
- Meyer M, Siniläinen T, Utecht J (2003) Towards hybrid triple helix indicators: a study of university-related patents and a survey of academic inventors. *Scientometrics* 58(2):321–350. <https://doi.org/10.1023/A:1026240727851>
- Mogoutov A, Kahane B (2007) Data search strategy for science and technology emergence: a scalable and evolutionary query for nanotechnology tracking. *Res Policy* 36(6):893–903. <https://doi.org/10.1016/j.respol.2007.02.005>
- Müller C (2002) The evolution of the biotechnology industry in Germany. *Trends Biotechnol* 20(7):287–290. [https://doi.org/10.1016/S0167-7799\(02\)01966-2](https://doi.org/10.1016/S0167-7799(02)01966-2)
- Murray F (2002) Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Res Policy* 31(8–9):1389–1403. [https://doi.org/10.1016/S0048-7333\(02\)00070-7](https://doi.org/10.1016/S0048-7333(02)00070-7)

- Noyons ECM, Luwel M, Moed HF (1998) Assessment of Flemish R&D in the field of information technology: a bibliometric evaluation based on publication and patent data, combined with OECD research input statistics. *Res Policy* 27(3):285–300. [https://doi.org/10.1016/S0048-7333\(98\)00046-8](https://doi.org/10.1016/S0048-7333(98)00046-8)
- Pammolli, F., & Rossi, M. A. (2005). Intellectual property, technological regimes and market dynamics. *Economia e Politica Industriale Paper*, (2). <https://ssrn.com/abstract=1031653>.
- Perry BL, Pescosolido BA, Borgatti SP (2018) Egocentric network analysis: foundations, methods, and models, vol 44. Cambridge University press, Cambridge
- Roesler C, Broekel T (2017) The role of universities in a network of subsidized R&D collaboration: the case of the biotechnology-industry in Germany. *Rev Regl Res* 37(2):135–160. <https://doi.org/10.1007/s10037-017-0118-7>
- Rueda G, Gersdri P, Kocaoglu DF (2007) Bibliometrics and social network analysis of the nanotechnology field. PICMET'07-2007 Portland international conference on management of engineering & technology. IEEE, Piscataway. <https://doi.org/10.1109/PICMET.2007.4349633>
- Schoen A, Heinisch D, Buenstorf G (2014) Playing the ‘name game’ to identify academic patents in Germany. *Scientometrics* 101(1):527–545. <https://doi.org/10.1007/s11192-014-1400-x>
- Silge J, Robinson D (2017) Text mining with R: A tidy approach. O'Reilly Media, Inc, Newton
- Simmie J (2002) Knowledge spillovers and reasons for the concentration of innovative SMEs. *Urban Stud* 39(5–6):885–902. <https://doi.org/10.1080/00420980220128363>
- Tahmooresnejad L, Beaudry C (2019) Collaboration or funding: lessons from a study of nanotechnology patenting in Canada and the United States. *J Technol Transf* 44(3):741–777. <https://doi.org/10.1007/s10961-017-9615-7>
- Tan J, Fu HZ, Ho YS (2014) A bibliometric analysis of research on proteomics in science citation index expanded. *Scientometrics* 98(2):1473–1490. <https://doi.org/10.1007/s11192-013-1125-2>
- Verbeek A, Debackere K, Luwel M, Andries P, Zimmermann E, Deleus F (2002) Linking science to technology: using bibliographic references in patents to build linkage schemes. *Scientometrics* 54(3): 399–420. <https://doi.org/10.1023/A:1016034516731>
- Wang B, Liu Y, Zhou Y, Wen Z (2018) Emerging nanogenerator technology in China: a review and forecast using integrating bibliometrics, patent analysis and technology roadmapping methods. *Nano Energy* 46: 322–330. <https://doi.org/10.1016/j.nanoen.2018.02.020>
- Wanzenböck, I., Schemgell, T., & Brenner, T. (2013) What determines the position of regions in European knowledge networks? A comparative perspective on R&D collaboration, co-patent and co-publication networks. <http://hdl.handle.net/10419/123909>.
- Wasserman S, Faust K (1994) *Social network analysis: methods and applications*, vol 8. Cambridge university press, Cambridge
- Zhang G, Liu L, Wei F (2019) Key nodes mining in the inventor–author knowledge diffusion network. *Scientometrics* 118(3):721–735. <https://doi.org/10.1007/s11192-019-03005-2>
- Zhu L, Zhu D, Wang X, Cunningham SW, Wang Z (2019) An integrated solution for detecting rising technology stars in co-inventor networks. *Scientometrics* 121(1):137–172. <https://doi.org/10.1007/s11192-019-03194-w>
- Zucker LG, Darby MR (1996) Star scientists and institutional transformation: patterns of invention and innovation in the formation of the biotechnology industry. *Proc Natl Acad Sci* 93(23):12709–12716. <https://doi.org/10.1073/pnas.93.23.12709>