ARTICLE IN PRESS

Technological Forecasting & Social Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Technological speciation as a source for emerging technologies. Using semantic patent analysis for the case of camera technology☆

Martin G. Moehrle*, Hüseyin Caferoglu

University of Bremen, IPMI - Institute of Project Management and Innovation, Wilhelm-Herbst-Strasse 12, D-28359 Bremen, Germany

ARTICLE INFO

Keywords:
Technological speciation
Emerging technology
Semantic patent analysis
Text-mining
Demand-pull
Camera technology management

ABSTRACT

In this paper we present a novel method which enables an early and direct detection of technologies emerging from a mainstream technology due to technological speciation. This method uses variables that were originally introduced to characterize emerging technologies such as novelty, persistence, growth, and community. It is applicable to mainstream technologies and relies mainly on semantic patent analysis. We test it in the field of camera technology, which has a longstanding tradition and has been influenced by several technological generations. Based on a patent search, we develop a process that comprises three steps, starting with the extraction and evaluation of bi-grams from the patents, continuing with the identification and evaluation of patents with novel and persistent bi-grams, and concluding with the identification of application fields and technological speciation candidates. As a result, we observe several instances of technological speciation, such as the action camera, the depth camera and the dashboard camera. Our approach involves theoretical, managerial, and political implications; for example, it helps companies establish a system for the early identification and monitoring of emerging technologies.

1. Introduction

When the carbon fiber reinforcement technology came into existence, it dramatically influenced the way in which airplanes are built; it also led to opportunities in other industries such as the automotive, bicycle and wind turbines industries (Moehrle and Passing, 2016). Carbon fiber reinforcements are a typical example of an emerging technology. The significance of an emerging technology is not only due to the new opportunities that it may offer directly, but also to a multitude of impacts that may originate from it indirectly. An emerging technology has "*the potential to create a new industry or transform an existing one*" (Day and Schoemaker, 2000) with an immense impact on economy, society, and politics (Hung and Chu, 2006; Martin, 1995; Porter et al., 2002). Therefore, there is continuous interest in the topic of emerging technologies on the side of researchers, policy-makers, and companies, which is reflected in the increase of publications related to emerging technologies over the last decades (Rotolo et al., 2015).

One important source of emerging technologies is the technological speciation from a mainstream technology (i.e. a mature, broadly elaborated technology covering several generations of development). In analogy to evolutionary biology, Adner and Levinthal (2002) define technological speciation as a phenomenon marked by two characteristics: First, an existing technology finds use in a new application domain, often represented by a specific niche of customer needs. Second, the technology is thereby often developed further in a way that differentiates the emerging technology from the lineage of the original technology.

There are many examples of how emerging technologies can be detected on the basis of patents. Almost all existing methods identify emerging technologies by focusing on aspects of scientific-push, namely the emergence of new keywords or classes of patent classification (e.g. Érdi et al., 2013; Joung and Kim, 2017; Kim et al., 2008; Yoon and Kim, 2011). They miss the opportunity to identify emerging technologies driven by demand-pull, as explained above.[1] Taking this opportunity into account, the question arises, whether it is possible to develop a method that enables an early and direct detection of technologies originating from technological speciation and qualifying them as emerging technologies.

In this paper we present a method of this sort. We take an existing mainstream technology as a basis, search for technological speciation

☆ This research was not supported by any specific grant from funding agencies in the public, commercial, or non-profit sectors.

* Corresponding author.
E-mail addresses: moehrle@uni-bremen.de (M.G. Moehrle), huseyin.caferoglu@innovation.uni-bremen.de (H. Caferoglu).

[1] Methods driven by scientific-push manage to identify technological speciation from existing technologies accidentally, when such technological speciation leads to a peak of patents granted in certain classifications or related to keywords from their search.

https://doi.org/10.1016/j.techfore.2018.07.049
Received 1 November 2017; Received in revised form 5 July 2018; Accepted 21 July 2018

candidates by semantic patent analysis and check for emerging technologies by applying a set of broadly accepted criteria. The paper is organized as follows: First, we provide the theoretical background. Second, we introduce two novel methods which are based on semantic analysis. Third, we select the camera industry as a promising case for testing the method and developing a related patent set. Fourth, we explain our method, focusing on patents especially and proceeding in three steps to identify so-called technological speciation candidates. Fifth, we present the results of our study in terms of the camera industry, identifying several technologies emerging from technological speciation, such as the action camera or depth camera technology. Sixth, we draw our conclusions. As a major point, we suggest the investigation of application fields as source for emerging technologies in addition to the traditional source of investigating scientific discoveries. In consequence, we suggest our method for technology managers to identify opportunities in new application fields.

2. Theoretical background

The research into emerging technology has captured the attention of numerous academic scholars. There are several approaches to the definition of this phenomenon (e.g. Boon and Moors, 2008; Martin, 1995; Porter et al., 2002; Small et al., 2014). A comprehensive definition is presented by Rotolo et al. (2015), who suggest a five-criteria model which unites and consolidates definitions by different researchers. According to them, an emerging technology is characterized by (i) radical novelty, (ii) relatively fast growth, (iii) coherence, (iv) prominent impact, and (v) uncertainty and ambiguity. While these criteria describe the nature of emerging technologies quite well, they are only partly accessible for operationalization (see the Method part of this paper). To give a theoretical background, we will outline reasons for the emergence of a technology; focus on demand-pull as driver for technological speciation and comment on the identification of emerging technologies within technological speciation.

2.1. Emergence of a technology

There may be different reasons for the emergence of a technology. Like other novel technologies, emerging technologies are driven by a scientific-push (which usually is labeled as *technology* push, but for distinguishing from emerging *technologies*, we use scientific-push), a demand-pull, or a combination of both (Dosi, 1982). For instance, it is well known that a breakthrough in natural science, such as the discovery of a new material like Graphene, represents a scientific-push and may not only lead to the emergence of a single technology, but a multitude of them. Such scientific-push offers new solutions for known tasks; at times it may even create new tasks. In contrast, a new market need, such as the increasing desire for clean air, may also foster the emergence of technologies for emission avoidance. Therefore, a demand-pull seeks for a new configuration or combination of known (or sometimes novel) technologies related to a new task. The combination of both – scientific-push and demand-pull – is what many emerging technologies are driven by. For instance, the satellite technology makes use of advanced knowledge about radio transmission and serves the people's need to communicate over long distances or determine their current location.

There is extensive theoretical and empirical work made for emerging technologies driven by scientific-push in any form (be it solely by scientific-push or by the combination of scientific-push and market-pull). Recent examples include organic LED by Shen et al. (2010), zinc oxide nanostructures by Ávila-Robinson and Miyazaki (2013), or thin film solar cells by Yoon et al. (2011). In contrast, there is a lack at least of empirical work for emerging technologies driven by demand-pull primarily.

2.2. Demand-pull as driver for technological speciation

A theoretical work in context of demand-pull driven technologies is proposed by Adner and Levinthal (2002). The authors investigate market needs and find that technological speciation can be a specific form of the emergence of technologies. In analogy to evolutionary biology, they coin the term technological speciation, describing it as a phenomenon in which an existing technology finds use in a new application domain, often represented by a specific niche of customer needs. In some cases, technological speciation may take place without significant changes of its technological antecedents (Adner and Levinthal, 2002). However, in many cases, significant changes are necessary; knowledge from other sources has to be integrated to fulfill the needs of the new application domain. As a consequence, technological speciation can lead to the emergence of new technologies which follow a different path than that of their technological antecedents (Adner and Levinthal, 2002). This view is further based on an evolutionary perception of a society's development (Alexander et al., 2012), which is characterized by a feedback loop: Changes in the environment or in cultural behavior lead to new or altered demands, while the satisfaction of these and earlier demands may lead to changes in the environment or in cultural behavior.

In the course of speciation, the technology adapts to the selection criteria, e.g. market needs, of the new application field in terms of functionality (Adner and Levinthal, 2002; Basalla, 1988). New functions may be added, and old functions may be changed or eliminated. In other words: a previously irrelevant function of a technology could be of relevance in its new field of application and may therefore be elaborated (Adner and Levinthal, 2002). Furthermore, resources are a crucial element of technological speciation. Only a combination of the adaptation process and the existence of resources in the new application field ensure the development of a new technology in a substantial way (Adner and Levinthal, 2002). Following this line of thinking, a new technology is developed on the basis of an existing technology, recombined with other existing technologies (Fleming, 2001). The new technology possesses characteristics which none of the recombined technologies used to have and it integrates a new kind of know-how to fulfill the changed requirements. As Garnsey et al. (2008) point out, such development combines both continuity and discontinuity - the continuity of the technology as a basis, and discontinuity of its phenomenology in its environment. It has to be noted that not every technology which comes into existence in this way is at the same time an emerging technology. In addition to its novelty, it is subject to the other aspects that emerging technologies are characterized by, such as relatively fast growth or coherence (Rotolo et al., 2015).

Our theoretical approach is an extension of Alexander et al. (2012), who characterize speciation of a technology "*as the point where a technology bifurcates between an existing technology and newer-generation of technology that at some point ends up competing against its 'parents'*" (Alexander et al., 2012). Our concept of emerging technologies caused by speciation is not limited to those competing with their 'parents'. Without neglecting this aspect, we also consider the set of speciation-caused technologies, which do not compete, to be relevant and rich.

2.3. Identification of emerging technologies within technological speciation

As in our understanding technological speciation is demand-pull driven, we find various opportunities to identify a basis for such emerging technologies (we will comment on the methods for extracting knowledge in the subsequent section). First, the researcher could focus on a specific application field and try to figure out whether it involves any emerging topics. Using patents as a basis, the researcher could refer to application-oriented patent classes or keywords to extract a basic set of patents, which can then be analyzed chronologically along with all incorporated technologies. For instance, if the researcher is interested in the catching or trapping of animals, CPC class A01M would be an appropriate source. Second, as speciation is usually based on existing technologies, a researcher could focus on an existing technology (one we might call a mainstream technology) with its full scope of applications. Again, the researcher could use technology-oriented patent classes or keywords to extract a basic set of patents, which can then be analyzed chronologically. For example, the CPC class F02 comprises patents regarding combustion engines. We consider focusing on a mainstream technology to be more promising, as it seems difficult to tell significant from insignificant information in a multitude of possible application fields.

To sum the theoretical background up: Not all emerging technologies result from speciation and not all speciation entities lead to emerging technologies. However, speciation can be seen as a remarkable driver for emerging technologies, which has not been focused on yet. Moreover, methods have to be developed, which identify speciation candidates and qualify them as emerging technologies.

3. Methodical background

Given the theoretical approach mentioned above – i.e. focusing on a mainstream technology, searching for technology speciation, and classifying some of the identified technologies as emerging technologies – a methodical concept is needed to define emerging technologies. For this purpose, we rely on the work by Newman and Suominen (2017) who present a consistent approach by defining emerging technologies by four attributes, namely novelty, persistence, growth and community (which is a slight variation of the definition of Rotolo et al. (2015), but better to operationalize).

Furthermore, there are a few established patent based methods that might be used to identify emerging technologies, although not without limitations. Let us briefly introduce two examples: Ranaei and Suominen (2017) and Joung and Kim (2017). Both were published recently, and are based on semantic analyses. They manage to overcome the deficits of bibliography based papers, namely the long time lag between the application for a patent and its classification as well as its citation.

Ranaei and Suominen (2017) present an unsupervised machine learning approach by applying topic modeling, i.e. latent Dirichlet allocation (LDA) and dynamic topic modeling (DTM), to a large number of vehicle related patent data in order to identify patterns of emergence. In the case of LDA, the authors manage to identify topics and, by adding the time aspect, observe the dynamism of topics, e.g. an increase or decrease of significance. In the case of DTM, the time aspect is taken into account from the very beginning, and thus topics are identified on a yearly basis. Consequently, it is possible to observe the dynamism of words, e.g. the increase or decrease of a word's significance in the context of a given topic.

While this approach certainly has its advantages, it also involves certain shortcomings due to methodical elements. First, the identified topics do not exclusively represent technologies. Hence, an ex-post qualitative analysis is required to clean the topics, and to assign topics or terms to technologies. Second, since LDA only delivers weak signals regarding the topic's significance, there is still need for further analysis regarding novelty, persistence, growth, and community in order to determine whether a topic or term is emergent. Third, DTM demands a lot of processing power and also is very time consuming. For this reason, the authors had to limit their analysis to a very small number of topics, i.e. three.

Joung and Kim (2017) suggest an informetric measure for monitoring emerging technologies and apply their method in a case of electron transfer mechanism in electrochemical glucose biosensors. The chief progress of this approach is that the authors work with technical keywords that are automatically detected by means of an extended tf-idf[2] function. Furthermore, relationships between the technical keywords (e.g. synonym, hypernym, hyponym) are calculated and assigned by a domain expert. On the basis of identified technical keywords, the authors construct a new keyword to patent matrix which they use to calculate the dissimilarity matrix. An agglomerative clustering is then carried out to form clusters. Then, emerging technologies are detected by examining the clusters for emerging technical terms that were determined by domain researches.

Although this approach is promising as well, it has several shortcomings: First, a high level of manual intervention is required in the assignment of relationships between technical keywords, the detection of emerging technical keywords and the examination of clusters. Second, this approach merely considers two attributes of emerging technologies, i.e. radical novelty and coherence. Since novel approaches argue that emerging technologies are characterized by novelty as well as by persistence, growth, and significant community, the obtained results call for further analysis.

In summary: Semantic methods seem to possess great potential for identifying emerging technologies. They overcome the deficits of bibliographical methods. Manual and IT-based effort is required in the established methods. Both methods fail to be optimized for the identification of emerging technologies caused by technological speciation. For this reason, we see a potential to develop an own method and we expect in the future benchmarking efforts between all approaches.

4. Data

For our analysis, we choose a case technology and retrieve a dataset consisting of patents (patent pool). We decide to select camera technology as a case example, as we expect technological speciation to occur in this field. There are many reasons for our selection of cameras: Camera technology has a long-standing tradition. The first device of this kind – a *camera obscura* - was used in the 19th century. Since then, camera technology has been exposed to various influences, which could especially be observed in recent years. Following the introduction of smartphones, there has been a slump in the sales of compact cameras (Statista, 2012). This trend puts immense pressure on companies in the camera industry in terms of competition and innovation. As a result, we assume that there is an increase in innovative efforts in the camera industry.

In addition, we choose patents as an information source for our analysis. Patents are widely accepted for this purpose, as they cover most of the world's documented technical knowledge (Alberts et al., 2017). In contrast to other technologies such as software, camera technology is very well represented in patents. We proceed by the following three steps: First, we develop a basic search for camera technology. Second, precision and recall are used to analyze the quality of this search. Third, we structure the patent pool according to time periods.

For our data, it is important to delineate the technology at hand structurally as well as time-wise. In our understanding, a camera comprises several necessary structural elements, such as lenses, objective, imaging sensor, storage medium, image processor, and energy supply. In order to keep our focus on the mainstream technology, we exclude combinations of elements in accordance with our understanding of the technology as whole. In order to delineate the camera technology time-wise, a time-restricted title search is initiated (Alberts et al., 2017). We select the United States Patent and Trademark Office (USPTO) Patent and Full-Text Image Database (PatFT) and perform a keyword search with the aim to find a broad range of camera related patents, not only patents related to classes of cameras from the CPC. The search string "TTL/camera AND APD/01/01/1995- > 12/31/2015 AND ISD/01/01/1995- > 06/06/2017 AND APT/1" produces 15.036 patents, which are then used for further analysis.[3]

[2] Tf-idf stands for term frequency - inverted document frequency and can be calculated in different ways (Salton and Buckley, 1988).

[3] Although our search string might look simple, it is the result of extensive testing. In addition to the keyword "camera", we used parts of cameras such as lens, objective, image sensor in combination. Doing so, we could enhance the recall, but at the same time, the precision decreased. We preferred to use the efficient point with the simple search string and a high precision, which is helpful for the following procedure.

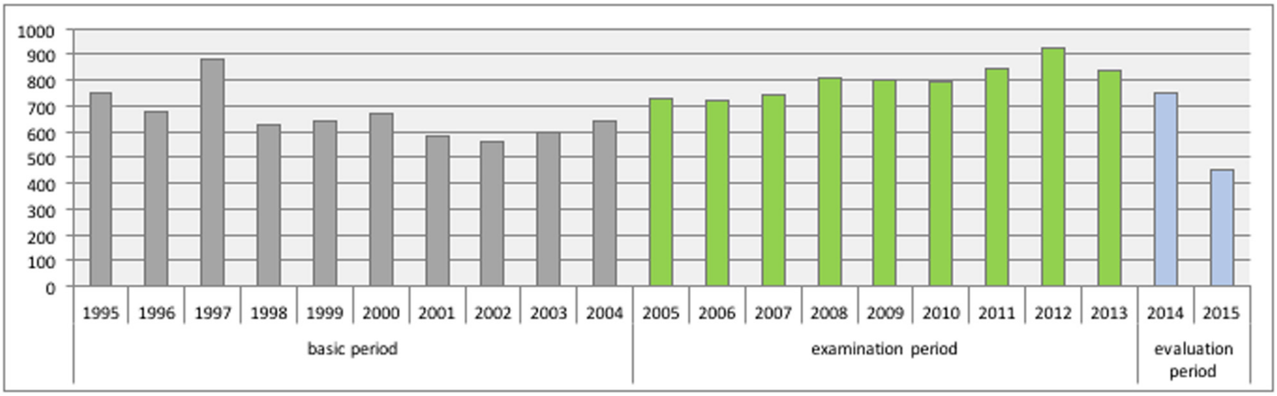

Fig. 1. Retrieved granted patents for mainstream technology per year, organized by application date.
Source: Authors.

For the evaluation of the search string's quality, (i) precision and (ii) recall are estimated following the suggestions given by Egghe (2008). (i) Precision is a measure that shows how many of the retrieved patents are relevant. We examine a sample of 50 patents that were selected randomly by means of an excel function, finding 36 thereof to be relevant. The other patents mainly describe inventions related to camera equipment and can at least be looked upon as partly relevant. We decide to accept this precision and forego a further cleaning of results. (ii) Recall is a measure which shows how many patents contained in the patent database as a main unit could be found by the search. We decide to expand the search, using additional keywords that are related to cameras or parts thereof. For this purpose, we extract the five most frequent Bi-grams in terms of document frequency and use them in a new title search, after eliminating duplicates. Doing so, 30.341 patents are identified as an upper bound, namely as potentially relevant hits. As the original search leads to a result of 15.036 patents, the recall amounts to at least 0,50.

Finally, the patent pool is divided into three periods in terms of application time, i.e. a basic period, an examination period, and an evaluation period. However, we actively search for technological speciation in the examination period only. The basic period and the evaluation period are needed for the calculation of certain measures that we require in order to determine a new technology's emergence. More concretely, a 20-year time frame extending from 1995 to 2015 is divided into three shorter periods (Fig. 1). The examination period is restricted to the years 2005 to 2013. The year 2005 deserves special attention as it is located in-between two significant phenomena. Firstly, the years prior to it were marked by the transition from analog to digital photography. Secondly, the iPhone was introduced in 2007. Thus, the sector was faced with new challenges, such as additional innovation pressure. As a result, we assume that the following years are marked by technological speciation. The years from 1995 to 2004 are selected as the basic time period; they are characterized by a digitalization of photography and a high number of patents. The evaluation period is restricted to only two years. As 2016 and 2017 are marked by a considerable decline in the publication of patents (in consequence of the patent granting process), the years 2014 and 2015 were chosen for this purpose.

5. Method

The aforementioned patent search yields the input data for our method to determine speciation candidates. Our approach adapts ideas and variables from the method devised by Newman and Suominen (2017) for the identification of emerging technologies. It comprises three steps (see Fig. 2): (i) extraction and evaluation of bi-grams per examination year, (ii) identification and evaluation of patents with novel and persistent bi-grams, and (iii) identification of application fields and speciation candidates.

According to Newman and Suominen (2017), the emergence of a technology is reflected by four variables, namely (i) novelty, (ii) persistence, (iii) community, and (iv) growth, which can be operationalized by means of different text-mining and bibliographic approaches.[4] (i) Novelty can be measured by the occurrence of new bi-grams in documents, in our case in patents. For the recognition of new bi-grams, we therefore recommend dividing the data (textual elements of patents) into time slices, in order to reveal the new bi-grams by a mutual comparison of those time slices. (ii) Persistence is related to the idea that the new bi-grams should be used to a certain degree after their first occurrence, in contrast to new bi-grams that occur only once in the entire period. (iii) An emerging technology is characterized by a community evolving around the topic, whose members maintain relationships (e.g. through citation) with one another (Newman and Suominen, 2017). It can be operationalized by measuring the number of patent applicants that appear in application-oriented classes of the CPC in a specific time frame. (iv) Growth reflects the increase of a topic's relevance (in our case a topic is defined as a patent with a significant number of new bi-grams). It can be measured by similarities between the topic and the remaining patent pool in time slices.[5]

5.1. Extraction and evaluation of bi-grams per examination year

The aim of the first step is to develop a bi-gram list per year for the examination period, based on the patent pool. We divide this step into four sub-steps, namely (i) selection of patent parts, (ii) language preprocessing, (iii) extraction of bi-grams, and (iv) novelty and persistence assessment.

(i) The first sub-step focusses on the selection of suitable textual parts of patents, which are the focal point of a semantic analysis. Based on recommendations by Tseng et al. (2007), Yoon et al. (2013),

[4] While we follow the authors' recommendations regarding the four variables, we operationalize them slightly differently.

[5] For the assurance of the last criterion, Newman and Suominen (2017) suggest developing a model, and simulating the procedure and the initiation process (growth) to potentially derive a forecast on this basis. It is therefore recommended to commence with gathering data regarding the frequency of bi-gram usage throughout time (Newman and Suominen, 2017). On the one hand, logistics curves are suitable for the growth evaluation (S-curves) (see Alexander et al., 2012; Young, 1993). On the other hand, probability models, such as the Latent Dirichlet Allocation, are suited for the tracking of time series as well as for forecasts (see Suominen et al., 2017; Suominen and Toivanen, 2016).

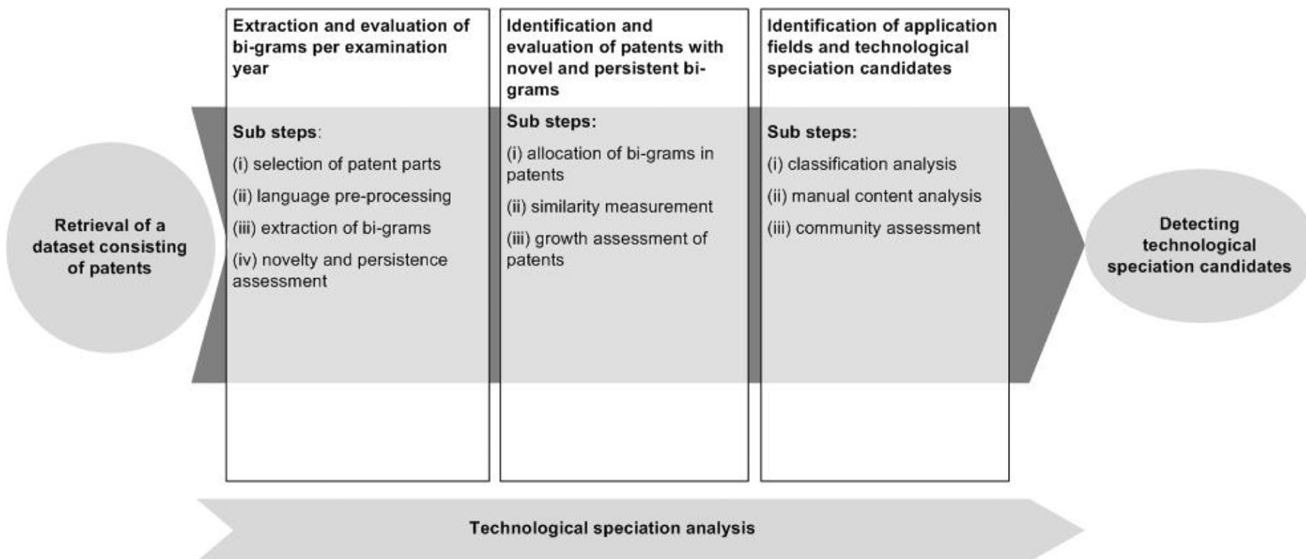

Fig. 2. Method for technological speciation analysis.
Source: Authors.

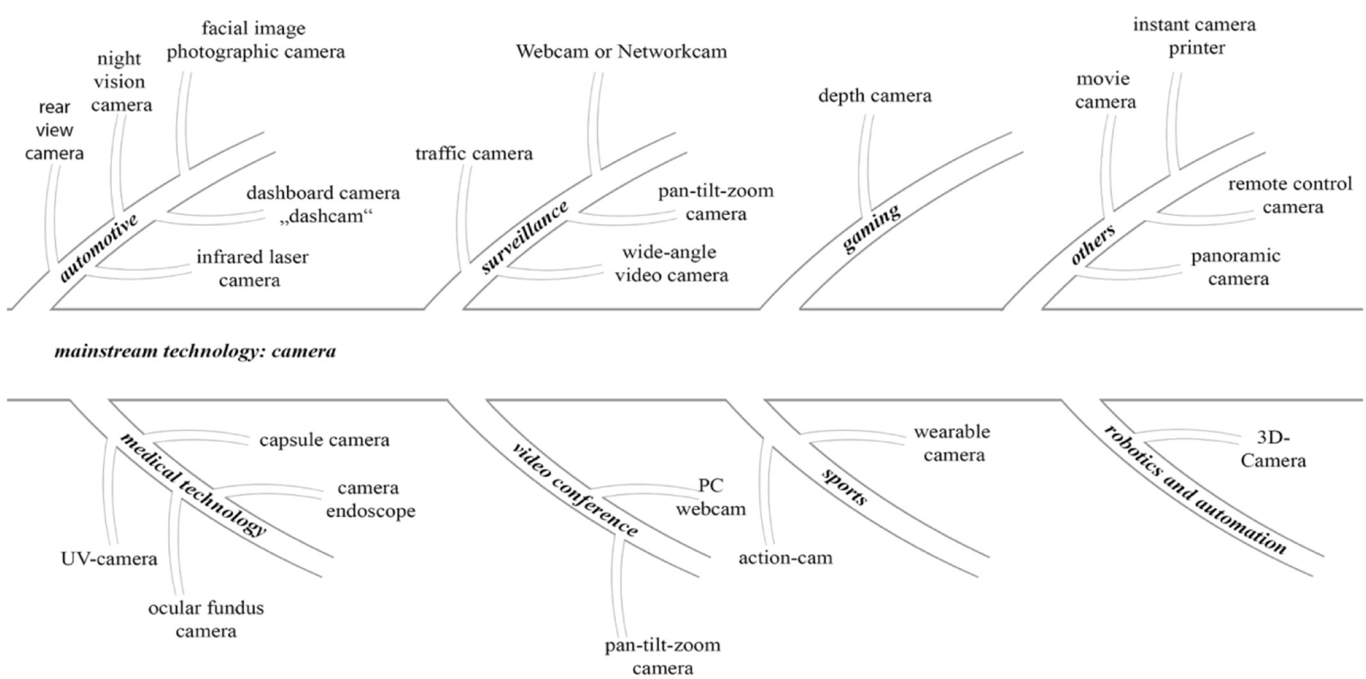

Fig. 3. Application fields and technological speciation candidates for camera technology. The application fields are organized according to the number of patents assigned to them, with one exception (application field “others” which is placed at the end).
Source: Authors.

Preschitschek et al. (2013), and Moehrle and Passing (2016), the title, abstract and claims are chosen, as this is where an invention's essential content is described.

(ii) By means of language pre-processing, a text can be cleaned (by filtering out stop words) and unified (by lemmatizing the terms). We use the PatVisor® which, for instance, eliminates stop words and patent specific words as well as numbers, punctuation and symbols (Walter et al., 2017). A lemmatizer is used to obtain word stems. Additionally, based on a uni-gram ($n = 1$) term-document matrix, bi-grams with a high document frequency are detected and eliminated if they fail to add contextual information. Terms of this type are: “method”, “apparatus”, “device”, “camera”, “system”, and “unit”.

(iii) Based on the recommendation given by Moehrle and Gerken (2012), the n-gram analysis is used for extracting semantic structures.[6] After inspecting the most frequent uni-grams, we set the

[6] We decide not to use uni-grams, as novel uni-grams are part of larger n-grams for which reason they are implicitly integrated in our method. Bi-grams get their novelty from two sources. They can contain novel uni-grams and/or they can represent novel combinations of uni-grams, which are not novel as such. The same is true for tri-grams, but in contrast to bi-grams the number of novel tri-grams increases dramatically, primarily due to novel combinations of uni-grams which are not novel as such. For this reason, and after testing, we decide to use bi-grams for our analysis.

Table 1
Application fields with their technological speciation candidates and relevant variables for emerging technologies.
Source: Authors.

Application fields	Technological speciation candidates	Novelty (represented by the identified bi-grams)	Persistence	Growth	Community	Top applicant (industry)	Search string for community
Automotive	Rear view camera; infrared laser camera; dashboard camera "dashcam"; facial image photographic camera; night vision camera	Configure laser; plate receiver; backup vehicle; windshield windshield; attach windshield; configure environment; portion windshield; compass sensor; collision vehicle; trajectory vehicle; algorithm vehicle; marker sensor; chronological image; maneuver vehicle; configure touchscreen	✓	✓	> 50	Magna Electronics, Inc. (automotive)	TTL/(camera) AND APD/01/01/2005- > 31/12/2013 AND CPC/(B60K$ or B60Q$ or B60R$ or B62D$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Medical technology	Camera endoscope; capsule camera; ocular fundus camera; UV-camera	Configure laser; instruction transitory; instruction transitory; capsule image; distal fluid; endoscope fluid	✓	✓	> 50	Canon Inc. (imaging and optical products)	TTL/(camera) AND APD/01/01/2005- > 31/12/2013 AND CPC/(A61F$ or A61B$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Surveillance	PTZ camera, Webcam or Networkcam, wide-angle video camera, traffic camera	Link scene; message notification; determine template; compare score	✓	✓	> 50	Panasonic Inc. (electronics)	TTL/(camera) AND APD/01/01/2005- > 31/12/2013 AND CPC/(G08B$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Video conference	PC-webcam and pan-tilt-zoom camera	Side socket; configure steer	✓	✓	> 50	Cisco Technology, Inc. (networking hardware)	TTL/("video and conference"~5 or "video call"~5 or) AND APD/01/01/2005- > 07/31/2013 AND CPC/(G03B$ or H04N$ or G02B$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Sports	Wearable camera; action-cam	Configure magnet; board sport; face fasten; enclose upper	✓	✓	10	GoPro, Inc. (imaging focus on sports)	TTL/(camera) AND SPEC/((sport or sports) AND (surfing or jogging or (mountain and climbing) or snowboarding or skydiving)) AND APD/01/01/2005- > 07/31/2013 AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Gaming	Depth camera	Render visual; feature mobile; transitory virtual; environment mobile, depth initial	✓	✓	17	Microsoft Inc. (gaming)	TTL/(camera) AND APD/01/01/2005- > 31/12/2013 AND CPC/(A63F$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1
Robotics and automation	3D-camera	Capture estimate; depth visual; filter system	✓	✓	21	InTouch Technologies, Inc. (healthcare)	TTL/(camera) AND APD/01/01/2005- > 31/12/2013 AND CPC/(B25J$ or A47L$ or E05F$ or A01J$) AND ISD/01/01/2005- > 06/06/2017 AND APT/1

minimal word length to 4, so that words of fewer than four letters are eliminated. A bi-gram analysis (i.e. $n = 2$) is performed, the window size is set to $n + 2$.

(iv) We now choose 100 bi-grams per analyzed year of the examination period[7] that fit both criteria, i.e. novelty and persistence. First, we check all bi-grams from all years of the examination period for novelty. We compare each bi-gram with all bi-grams that occurred earlier (also in the basic period) and flag the novel ones (growing time-frame technique). Only these are considered further. Now we check the persistence of novel bi-grams on a yearly basis. For this purpose, we suggest using the persistence frequency. Persistence frequency refers to the number of consecutive years in which a bi-gram was mentioned, starting with the year of examination. Doing this enables us to characterize each bi-gram in terms of its persistence frequency, and sort all bi-grams according to this variable in decreasing order. For each year, we pick 100 bi-grams based on the highest persistence frequency. If there are more bi-grams with the same persistence frequency than required, we additionally apply term frequency.

5.2. Identification and evaluation of patents with novelty per examination year

In the second step, patents with novelty are identified on the basis of the bi-gram lists. We proceed according to the following three sub-steps: (i) allocation of bi-grams in patents, (ii) similarity measurement, and (iii) growth assessment of patents.

(i) In the first sub step, we assign the extracted bi-grams of each examined year to the same year's patents. This step is aimed at identifying patents whose semantic structure incorporates one or more novel and persistent bi-grams.

(ii) Regarding each of the patents identified in the previous sub-step, we measure the semantic similarity between this particular patent and all other subsequent patents of the examination and test period individually with the aid of PatVisor®. Based on the recommendations given by Moehrle (2010), we choose Complete Linkage B and C as a variable measure and Double-Single-Sided-Jaccard for the similarity coefficient in the case at hand. As a result, PatVisor® produces a similarity matrix which is exported to Microsoft Excel. The patents from sub-step one form the rows, the subsequent patents form the columns of this matrix, in which the fields represent individual similarities. We now integrate the similarity values on a yearly basis by calculating the arithmetic average and adding respective columns to the matrix.

(iii) In the third sub-step, we identify patents with growing impact. We measure this impact by means of the relationship of two similarity values. The first of these two is the similarity value between the patent at hand and the set of patents from the same application year. The second is the arithmetic average of the similarity values between the patent at hand and the set of patents from all subsequent application years. If the second similarity value exceeds the first, the patent at hand has a greater similarity with patents from the examination and test period than from its application year. In that case, we classify the respective patent as one of growing impact.

5.3. Identification of application fields and speciation candidates

In the final step of the analysis, the previously identified patents with novel and persistent bi-grams and growing impact are scanned to identify technological speciation candidates. Again, we divide this procedure into three sub-steps, i.e. (i) classification analysis, (ii) manual content analysis and (iii) community analysis.

First, we identify application-oriented main classes of the CPC on the 4-digit level (in contrast to function-oriented main classes). We decide to use the CPC due to the fact that it involves a higher proportion of application-oriented classes than the IPC (Walter and Schnittker, 2016). We analyze each patent's main class and flag it, if the main class is application-oriented.

In the following sub-step all flagged patents are manually examined to establish their fields of application (which may differ from the CPC classification) and the role they might play with regard to upcoming products. The aim of this sub-step is to identify new application fields in which technological speciation might be expected, and cluster the related patents. If we find a product, which we see as manifestation of a technology, in this type of application field, we characterize it as being a technological speciation candidate. To achieve this, we examine each patent, in particular its title, summary, images and description, and try to assign the patent to a product.

In the final sub step, the application fields are submitted to a bibliographical analysis in order to answer the following question: Does a community evolve around a field of application, or is the field dominated by individual companies? To answer this question, we develop a separate search string for each application field. We use the initial search string as a baseline, adding the patent class, if it directly represents the application field. However, if the patent class does not describe the identified application field, the initial search string is complemented by further keywords that characterize the field of application (Alberts et al., 2017). We use a threshold value of 10 applicants. If an application field exceeds this threshold value, we assume it to involve an active community.

6. Results

Our analysis points to a variety of application fields as well as technological speciation candidates. By means of our method we manage to detect the following eight application fields: (i) automotive, (ii) medical technology, (iii) surveillance, (iv) video conference, (v) sports, (vi) gaming, (vii) robotics and automation, and (viii) others (see Fig. 3). (i) In the context of vehicles, we identify such technological speciation candidates as the rear view camera, infrared laser camera, dashboard camera "dashcam", facial image photographic camera, and night vision camera. (ii) Applied to the medical industry, we find the camera endoscope, capsule camera, ocular fundus camera, and UV-camera. (iii) The pan-tilt-zoom camera, webcam or networkcam, wide-angle video camera, and traffic camera are used in the context of surveillance. (iv) The PC-webcam and pan-tilt-zoom camera are used for video conferences and are thus regarded as further technological speciation candidates. (v) For use in outdoor sport activities such as skydiving or mountaineering we find the wearable camera and the action camera. (vi) Furthermore, we identify the depth camera, which is used in connection with game consoles. (vii) Another technological speciation candidate is represented by the 3D-Camera, which we assign to the application field of robotics and automation. (viii) Finally, we identify several patents that relate to camera equipment. Although these patents do not represent any specific application field, some of them point to further technological speciation candidates, such as the panoramic camera, movie camera, instant camera printer, or remote control camera.

The identified application fields match the variables suggested by Newman and Suominen (2017), albeit with a slightly different operationalization than described in their paper. In Table 1, we present the application fields and the related variables. For the illustration of novelty, we include the bi-grams that relate to the application fields. For instance, the novelty of technological speciation candidates in the application field of automotive is headed by bi-grams like configure laser or trajectory vehicle. The persistence of the bi-grams was evaluated as

[7] Due to the limited capacity of Excel, we had to restrict the number of bi-grams under inspection.

described above; the same applies to the growth of patents. To analyze the factor of community, we perform a bibliographic analysis focusing on applicants from 2005 to 2013. This, for instance, shows that more than 50 companies or individuals applied for patents in the application field of automotive technology.

In addition to the quantitative analysis, we identify the top applicant of each application field (see column “top applicant (industry)” in Table 1). In most application fields, companies belonging to this field head the list of applicants. For instance, Magna Electronics is an automotive supplier, Cisco Technology provides network hardware, and Microsoft is not only known as a provider of software for personal computers but also as one of the world's major gaming companies. Obviously, the traditional companies that belong to the mainstream technology allow room for entrants from other industries, leaving them opportunities to serve needs in specific fields.

In summary, technological speciation can be seen as a cause for the emergence of technologies. In our case, we base the analysis on the mainstream technology of cameras. Various new technologies, which we call speciation candidates, have been developed, and they all solve unique customer demands, fitting into a market niche, and incorporating and combining different sources of external knowledge. Some sources of external knowledge are novel, such as the 3D visual analysis software, some are well-established e.g. infrared laser technology. Consequently, the search for emerging technologies based on technological speciation may lead to results that significantly add to the classical scientific-push-driven search.

7. Conclusions

In this paper, we focus on technological speciation from a mainstream technology as a new source for detecting emerging technologies. We present a method based on the four variables suggested by Newman and Suominen (2017) to determine patents that reflect possible application fields and, integrated in these application fields, technological speciation candidates. In contrast to previous methods, we operationalize the variables in a different way, e.g. using bi-grams for a receiving a better embeddedness of terms and using a growing timeframe technique for the identification of novel bi-grams. By applying the method to the selected field of camera technology, we obtain insights regarding the evolution of the camera. As a result of our method, we identify several technological speciation candidates that emanate from the mainstream of camera technology, such as the action camera, dashboard camera, or depth camera. The method can easily be adapted to other fields of interest.

From a theoretical point of view, we contribute to the theory of emerging technologies, shedding light on technological speciation, the identification of related technologies and their qualification as emerging technologies. In contrast to scientific-push based methods, we find emerging technologies that were developed by combining existing technologies in a new way, such as the action-cam, which combines traditional camera technology, image stabilization software, water and shockproof housing. For this reason, our type of analysis may add new candidates of emerging technologies to the scientific-push driven type of analysis. Remarkably, it is possible to focus on a mainstream technology in order to identify technological speciation candidates. In general, our findings illuminate the importance of demand-pull driven technologies, opening up the field of search to application areas that are influenced by a broad spectrum of factors. In consequence, the search for emerging technologies is no longer limited to scientifically inspired technology experts. It is open to technology experts from mainstream technologies as well as experts from the application fields, looking for upcoming niches, be they sociologists seeking new forms of work force organization or marketing experts analyzing new ways of communication among young people.

From a managerial point of view, our method opens up paths in three directions. First, we take on the perspective of a company that is active in a mainstream technology (e.g. camera technology, in our case). This company could employ our method as an instrument for monitoring and scanning in order to observe technological speciation and identify developments in new application fields. This knowledge can be helpful, if the company is aiming for an extension of its product portfolio, but it may also be critical, if the identified technology speciation has the potential to substitute the companies' technologies. Second, we take on the perspective of a company that is active in a specific application field, but not yet in the respective mainstream technology. By application of our method, this company is enabled to detect a possible entrance to the mainstream technology and react accordingly. Third, as a strategic orientation instrument the method is particularly useful for companies that are looking for new forms of investment through the identification of new application fields.

From a political point of view, the results of our method can be used for the strategic alignment of national R&D. Government agencies could monitor technological speciation on an international level. If they discover interesting opportunities, they may change national R&D programs accordingly, for instance by investing in a new application field and dropping current developments that fail to offer similarly promising chances. Or they could stimulate international cooperation to obtain opportunities for national research and industry.

Our study is characterized by several limitations. (i) The method is limited to technological fields in which a patent search yields a high precision. If the precision of a patent search were low, a high quantity of non-relevant speciation candidates would result. (ii) Another limitation is caused by our search string, which is restricted to the word “camera”. There may be terms that occur over time and refer to a camera without incorporating the exact word, e.g. imaging device. This does not only limit our recall - if the effect is distributed unevenly over time, this may lead to misperceptions concerning the significance of specific developments. (iii) The use of arithmetic averages for growth assessment is also a limiting factor, as the arithmetic average is prone to outliers. Hence, a patent will be categorized as having a growing impact on the basis of values that are noticeably above average due to a “hype-phase” in one particular year. (iv) An analysis in terms of CPC main classes leads to a disregard of patents with function-oriented main classes, which may still be of an application-oriented character. (v) We examine patents manually to identify new application fields and technological speciation candidates therein. Although new products are easy to identify (for instance by combining the description and figures of a patent), there may be variances in their clustering, if done by different analysts.

Our work points out possibilities for further research. (i) The introduced method could be applied to other technological fields in which a mainstream technology exists. (ii) Subject-Action-Object structures could be used instead of *n*-grams for the extraction of semantic structures as well as the similarity analysis. (iii) For the semantic analysis, we use specific parts of patents, such as title, abstract and claims. Subsequent work could also refer to the descriptions, as these may contain information regarding the application fields. Unfortunately, there is no clear regulation for the description of a patent. For instance, the inventor might outline prior art extensively or explain the advantages of his invention compared to others. For this reason, it is not easy to identify the relevant information about an application field in the description. However, we regard this to be a possible topic for further research. (iv) Based on sophisticated diagram analysis, design patents could also be used as a means of identifying technological speciation candidates.

References

Adner, R., Levinthal, D., 2002. The emergence of emerging technologies. Calif. Manag. Rev. 45, 50–66.

Alberts, D., Yang, C.B., Fobare-Deponio, D., Koubek, K., Robins, S., Rodgers, M., Simmons, E., Demarco, D., 2017. Introduction to patent searching. In: Lupu, M.,

Mayer, K., Kando, N., Trippe, A.J. (Eds.), Current Challenges in Patent Information Retrieval. Springer Berlin Heidelberg, Berlin, Heidelberg.
Alexander, J., Chase, J., Newman, N., Porter, A., Roessner, J.D., 2012. Emergence as a conceptual framework for understanding scientific and technological progress. In: 2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies.
Ávila-Robinson, A., Miyazaki, K., 2013. Evolutionary paths of change of emerging nanotechnological innovation systems: the case of ZnO nanostructures. Scientometrics 95, 829–849.
Basalla, G., 1988. The Evolution of Technology. Cambridge Univ. Press, Cambridge.
Boon, W., Moors, E., 2008. Exploring emerging technologies using metaphors – a study of orphan drugs and pharmacogenomics. Soc. Sci. Med. 66, 1915–1927.
Day, G.S., Schoemaker, P.J.H., 2000. Avoiding the pitfalls of emerging technologies. Calif. Manag. Rev. 42, 8–33.
Dosi, G., 1982. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. Res. Policy 11, 147–162.
Egghe, L., 2008. The measures precision, recall, fallout and miss as a function of the number of retrieved documents and their mutual interrelations. Inf. Process. Manag. 44, 856–876.
Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., Zalányi, L., 2013. Prediction of emerging technologies based on analysis of the US patent citation network. Scientometrics 95, 225–242.
Fleming, L., 2001. Recombinant uncertainty in technological search. Manag. Sci. 47, 117–132.
Garnsey, E., Lorenzoni, G., Ferriani, S., 2008. Speciation through entrepreneurial spin-off: the Acorn-ARM story. Res. Policy 37, 210–224.
Hung, S.-C., Chu, Y.-Y., 2006. Stimulating new industries from emerging technologies: challenges for the public sector. Technovation 26, 104–110.
Joung, J., Kim, K., 2017. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. Technol. Forecast. Soc. Chang. 114, 281–292.
Kim, Y.G., Suh, J.H., Park, S.C., 2008. Visualization of patent analysis for emerging technology. Expert Syst. Appl. 34, 1804–1812.
Martin, B.R., 1995. Foresight in science and technology. Tech. Anal. Strat. Manag. 7, 139–168.
Moehrle, M.G., 2010. Measures for textual patent similarities: a guided way to select appropriate approaches. Scientometrics 85, 95–109.
Moehrle, M.G., Gerken, J.M., 2012. Measuring textual patent similarity on the basis of combined concepts: design decisions and their consequences. Scientometrics 91, 805–826.
Moehrle, M.G., Passing, F., 2016. Applying an anchor based patent mapping approach: basic conception and the case of carbon fiber reinforcements. World Patent Inf. 45C.
Newman, N.C., Suominen, A., 2017. Exploring the fundamental conceptual units of technical emergence. In: Journal Name: Proceedings of PICMET '17: Technology Management for Interconnected World.
Porter, A.L., Roessner, J.D., Jin, X.-Y., Newman, N.C., 2002. Measuring national 'emerging technology' capabilities. Sci. Public Policy 29, 189–200.
Preschitschek, N., Niemann, H., Leker, J., Moehrle, M.G., 2013. Anticipating industry convergence: semantic analyses vs IPC co-classification analyses of patents. Foresight 15, 446–464.
Ranaei, S., Suominen, A., 2017. Using machine learning approaches to identify emergence: case of vehicle related patent data. In: 2017 Portland International Conference on Management of Engineering and Technology (PICMET).
Rotolo, D., Hicks, D., Martin, B.R., 2015. What is an emerging technology? Res. Policy 44, 1827–1843.
Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. Inf. Process. Manag. 24, 513–523.
Shen, Y.-C., Chang, S.-H., Lin, G.T.R., Yu, H.-C., 2010. A hybrid selection model for emerging technology. Technol. Forecast. Soc. Chang. 77, 151–166.
Small, H., Boyack, K.W., Klavans, R., 2014. Identifying emerging topics in science and technology. Res. Policy 43, 1450–1467.
Statista, 2012. Are Smartphones Killing Digital Cameras? URL. https://www.statista.com/chart/5782/digital-camera-shipments/, Accessed date: 10 April 2017.
Suominen, A., Toivanen, H., 2016. Map of science with topic modeling: comparison of unsupervised learning and human-assigned subject classification. J. Assoc. Inf. Sci. Technol. 67, 2464–2476.
Suominen, A., Toivanen, H., Seppänen, M., 2017. Firms' knowledge profiles: mapping patent data with unsupervised learning. Technol. Forecast. Soc. Chang. 115, 131–142.
Tseng, Y.-H., Lin, C.-J., Lin, Y.-I., 2007. Text mining techniques for patent analysis. Inf. Process. Manag. 43, 1216–1247.
Walter, L., Schnittker, F.C., 2016. Patent Management. De Gruyter Oldenbourg, Berlin.
Walter, L., Radauer, A., Moehrle, M.G., 2017. The beauty of brimstone butterfly: novelty of patents identified by near environment analysis based on text mining. Scientometrics 111, 103–115.
Yoon, J., Kim, K., 2011. Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks. Scientometrics 88, 213–228.
Yoon, J., Choi, S., Kim, K., 2011. Invention property-function network analysis of patents: a case of silicon-based thin film solar cells. Scientometrics 86, 687–703.
Yoon, J., Park, H., Kim, K., 2013. Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis. Scientometrics 94, 313–331.
Young, P., 1993. Technological growth curves. Technol. Forecast. Soc. Chang. 44, 375–389.

Martin G. Moehrle has been Director of the Institute for Project Management and Innovation (IPMI) at the University of Bremen since 2001 while also holding the chair of Innovation and Competence Transfer. His area of work is technology and innovation management, and his preferred topics are patent management, technology road-mapping, future research, innovation process management, and methodical invention (TRIZ). He obtained his doctorate and qualified as a university lecturer on transitional themes between technology management and business informatics.

Hüseyin Caferoglu is a research associate at the Institute of Project Management and Innovation (IPMI) at the University of Bremen since November 2017. His research interests comprise technology and innovation management, intellectual property management, and semantic patent analysis. He holds an M. Sc. degree in Industrial Engineering and Management from the University of Bremen.