Hitotsubashi University
Institute of Innovation Research

The determinants of parallel invention: Measuring the role of information sharing and personal interaction between inventors

Byeongwoo Kang
Rudi Bekkers

IIR Working paper WP#22-06

2022年11月

一橋大学イノベーション研究センター

東京都国立市中2-1
http://www.iir.hit-u.ac.jp

The determinants of parallel invention: Measuring the role of information sharing and personal interaction between inventors

[Draft, November 1st, 2022]

Byeongwoo Kang* and Rudi Bekkers**

* byeongwoo.kang@iir.hit-u.ac.jp
Institute of Innovation Research, Hitotsubashi University, Japan

** r.n.a.bekkers@tue.nl
School of Innovation Sciences, Eindhoven University of Technology, The Netherlands

Historical accounts describe numerous cases of parallel invention. Nowadays, with over half a million inventions yearly that apply for patent protection at the USPTO alone, it is likely that there are a lot of parallel inventions among these. Yet, the mechanisms behind creating similar knowledge remain unstudied. From both a theoretical and practical perspective, it is an interesting question to what degree parallel inventions take place truly independent of each other, or whether they are the result of the exchange of knowledge and ideas between inventors. In our empirical study, we use the unique setting of technical standardization, where it is possible to systematically observe knowledge sharing as well as knowledge exchanges between inventors in detail. This study presents two novel analyses, one focussing on the determinants of similar inventions (using an AI-based approach) and one on the determinants of identical inventions (exploiting data from the patent granting procedure). In both analyses, we find positive and significant effects for knowledge sharing as well as for inventor interaction as determinants. The latter effect is the strongest: if meet in person and discuss their ideas, the likelihood of similar inventions increases up to a factor of approximately five, to up to 2.3 percentage points. Empirically confirming the theoretical work of Amabile (1983, 1988) on knowledge creation at the individual level and that of Nonaka (1994, 2006) on knowledge creation at the organizational level, we reflect on the implications of our findings for companies wishing to increase their inventive efforts.

Keywords:
Creativity; Idea twin; Knowledge creation; Patent similarity; Similar knowledge

JEL classification:
O31, D82

Disclaimer / funding
This work was supported by JSPS KAKENHI Grant Numbers 19K13768 and 20KK0296.

Acknowledgements
We would like to thank Atsushi Ohyama for providing constructive comments.

1 Introduction

The development of new ideas, knowledge or inventions is highly valued in society. If it is only a matter of time for a certain idea to come up (as suggested by Zeitgeist theories; see below), then it will also not come as a surprise that the same idea can come up more than once, completely independent of each other, either at about the same time or with some time interval. The phenomenon of parallel invention, also known as multiple discovery or simultaneous invention, has received considerable attention in the literature (see Simonton, 1979). Historical cases include the discovery of calculus by Newton but also by Leibniz (Khamara, 2006; Bardi, 2009), the theory of evolution of species, by Charles Darwin but also by Alfred Russel Wallace (Randy, 2008). Other examples are technical inventions such as the crossbow, claimed to be invented in both Europe and China (Foley et al., 1985) and the invention of the telephone by Alexander Graham Bell but also by Elisha Gray. But identical ideas might also not be totally independent of each other. Their likelihood may increase with similar training, with being exposed to the same set of knowledge or the same challenge to be solved, or with interaction and communication between prospective inventors. Or ideas may even have been copied. In fact, several of the above examples of independent discovery were or are contested, such as the telephone (see Brooks, 1975, who contests independent parallel invention of the telephone, as well as Shulman, 2009; Evenson, 2000, and Baker, 2000).

In current days, the incentives and expectations for researchers and engineers to come up with new ideas are high. At the USPTO alone, more than half a million inventions are filed for patent protection yearly (USPTO, 2020). This likely includes a large number of parallel inventions. At the same time, our understanding of knowledge creation and the potentially resulting parallel inventions is limited. Why do inventors come up with the same idea? What role do background/training, concurrent knowledge and personal interaction between inventors play in the invention process? This paper aims to contribute to such an understanding by an empirical study and uses a specific context allowing us to collect data that can disentangle several dimensions important to understand parallel invention. More specifically, we address the following research question: Assuming similar training and background, to what degree can parallel invention be explained by (1) inventors being exposure to the same (common) set of information on technical challenges/goals to be achieved or (2) the presence of personal interaction between inventors?

Our empirical data comes from the context of technical standard-setting, and, more specifically from the development of standards for mobile telecommunications in the 3GPP consortium. We collected data in this specific context because it offers a unique opportunity to observe the moments at which engineers get access to the same set of knowledge (including that of new challenges ahead that need to be solved) as well as if and when those engineers come together in the same space and discuss. The specific 3GPP setting is not only suitable in that it allows us to allow (separate) observation for these two types of events, but also because of its detailed paper trail on these events, such as full meeting details, including which participants were actually present. Furthermore, our empirical data uses patents as the result of knowledge development – and to measure parallel inventions. In the context of 3GPP, the strong incentives for companies to obtain Standard Essential Patents (SEPs) means that almost all knowledge developed is appropriate via patenting. Following two alternative ways to measure the occurrence of parallel invention – one relying on an AI-based approach, the other on assessments done by patent examiners, our study empirically concludes that when individuals and teams are exposed to identical knowledge input, they are likely to develop new similar knowledge.

Our study contributes to management and economics literature that focuses on knowledge development. While the phenomenon of parallel invention and the underlying theory has already received considerable attention, the current literature lacks empirical testing. This paper, to the best of our knowledge, is the first attempt at a systematic, quantitative study of parallel inventions. Our findings strengthen existing theories on knowledge development and creativity. Furthermore, our findings imply that exposure to new information plays an important role in creating new knowledge. To encourage innovation, both on the individual and the organisational level, one organisations may create an environment that promotes exposure to new information.

The remainder of this paper is structured as follows. Section 2 reviews the existing literature and develops hypotheses. Section 3 explains why the context of standardisation presents unique opportunities to empirically investigate parallel inventions. Section 4 introduces our empirical model, and Section 5 provides information on our dataset and main variables. The next two sections present our main analyses, first for similar inventions (Section 6) and then for identical inventions (Section 7). Finally, Section 8 concludes, discusses research implications, and points out the limitations of this study.

2 Literature review and hypothesis development

In this section, we review the literature on knowledge creation and the underlying processes, paying special attention to studies that model knowledge creation. Aiming to empirically test the predictions in that literature, we then develop working hypotheses.

2.1 Literature review: Modelling knowledge creation

Theories on knowledge creation have been developed both at an individual level and at an organizational level. For knowledge creation at an individual level, we take as a starting point the highly cited work by Teresa M. Amabile (1983, 1988), which was discussed and extended by subsequent studies (Farr & Ford, 1990; Kim, 2017; Groenewoudt et al., 2019). For knowledge creation at an organizational level, we start from the work of Ikujiro Nonaka (1994), which has been followed by many high-impact studies (e.g., Nonaka & Konno, 1998 and Nonaka et al., 2006).

The knowledge creation models in both these strains of literature contain similar elements. In the core, knowledge creation is modelled by the simple equation $y = f(x)$, where new knowledge (y) is a function of information input (x). Information input includes scientific knowledge as well as non-scientific knowledge such as mere information, recognized problems, communications, and senses. They stimulate an individual's brain to create new knowledge.

The knowledge creation function is formed by many factors. An individual's function depends on formal and informal training, personal experiences, and personality characteristics (Amabile, 1983; 1988). The function defines ways to perceive information input. Based on the function, an individual analyzes information input, decomposes it into elements, and finds responses. Each individual forms a different function of knowledge creation. Since each individual is not free from his/her unique social, cultural, and historical contexts (Vigotsky, 1986; Nonaka and Takeuchi, 1995; Jehn et al., 1999), knowledge creation function is bound by those personal contexts

However, when one deals with technical knowledge creation or with expertise creation, knowledge acquired from education and training is arguably much more important than personal contexts (Farr & Ford, 1990; Oldham & Cummings, 1996; Hammond et al., 2011). After all, to create new wireless communication technologies, one must have acquired knowledge in electrical engineering, electronics engineering, information processing, etc. In the same manner, to create new medical knowledge, one must have acquired knowledge in biology, chemistry, anatomy, physiology, etc. One decomposes and analyses information input based on the knowledge creation function formed through formal education. If information input involves technical challenges, then new technical solutions are created. If information input is scientific curiosity, then new scientific theories are created.

2.2 Hypothesis development

Considering the above literature, for our study, we will assume that individuals who have received education and training in the same field are likely to have very similar information processing mechanism. Accordingly, when they are exposed to an identical information input, there is a likelihood that they create the same knowledge. For our study, we will focus specifically on the field of engineering and on the role of information exchanges in the knowledge development process, assuming that the background knowledge of the involved individuals (i.e., education and training) will be similar. We will also focus on knowledge development on the organisational level, and more specifically on the firm level, and study not between engineers within the same firm (who often work together on a single invention). Below, we presume a setting where there are meetings taking place between companies, where participants (firm representatives) are subject to knowledge exposure, and where also personal communication between participants can take place.

Considering the theories discussed above, we formulate the following hypothesis concerning knowledge exposure:

H1: If firms are exposed at a given meeting to the same set of knowledge about challenges and goals, the likelihood increases that after that meeting, these firms will bring forth more parallel inventions.

Second, we focus on the role of personal interaction between individuals of different firms. Building on Nonaka & Konno (1998), we expect that when engineers discuss and interact in the same space, sharing context and tacit knowledge, the likelihood increases that they create similar knowledge and, thus, bring forward parallel inventions. Because of intensive knowledge exchange between engineers, we assume that when members are present in the same space, more similar knowledge is created than when members have only shared knowledge input without being in the same space. Since this is about personal interaction, we consider the individuals (as participants of meetings, but also as inventors). We formulate the following hypothesis:

H2: If individuals of different firms are personally participating at those given meetings, the likelihood increases that after that meeting, these individuals will bring forth more parallel inventions.

It is important to note that the second hypothesis is complementary to the first; the situation in the second hypothesis already implies presence at the same meetings, (and thus implies common exposure to knowledge). We further note that, out of these meetings, in the idle time, personal interaction may take place between individuals of the same company, but we leave that out of the scope of our analysis (and do not consider parallel inventions by individuals working for the same firm).

3 Standardisation as a context to understand parallel inventions

Given our hypothesis, the major challenge in collecting data is to be able to observe, in a systematic manner, (1) when they are exposed to a common set of information, e.g., on technical challenges/goals, (2) personal interaction between prospective inventors, and (3) the relevant inventions they bring forth, which may or may not be the result of that knowledge exchange. We carry out our data collection in the specific context of technical standardisation at 3GPP, a collaboration between the major regional standard-setting organisations in the field of telecommunications. The standardisation process is a knowledge creation process (Abdelkafi et al., 2021), and, in 3GPP, engineers invent new technologies to provide faster, more reliable communications technologies, catering for an increasingly wide range of applications and usages. With over 5.3 billion people across the globe that use mobile devices based on 3GPP standards (such as the 3G W-CDMA and 4G LTE standards), and another 15 billion Internet-of-Things (IoT) devices using these standards (GSMA, 2022:4), the 3GPP standards are among the most successful technical standards ever. Yet, we did not select 3GPP for the success of its standards but because of its standard development process and its associated publicly available paper trail, which makes it uniquely suitable to test our hypothesis on knowledge development. Below, we briefly sketch why this 3GPP context is uniquely suitable to answer our research questions; for a more in-depth discussion of 3GPP itself, we refer to Bekkers, 2001 and Kang & Bekkers, 2015.

The first aspect that makes 3GPP attractive for our case is that it carries out technical standardization. Employing the principles Open Standardization as laid out by the World Trade Organization (WTO, 2000), this means, among other things, that there is a transparent process in which technical development takes place, which is extensively documented. Consequently, we can observe knowledge exchange in a very detailed way, whereas normally, knowledge exchange that takes place within and across companies is not publicly available or otherwise observable in a systematic manner. The 3GPP context allows us to observe both knowledge sharing and personal inventor interaction:

- Knowledge sharing takes place via the documents and minutes that are available right after a 3GPP meeting to all participants. This way, the members (at the firm level) obtain the same information about the new functionalities and goals to be achieved, and can start working on this,

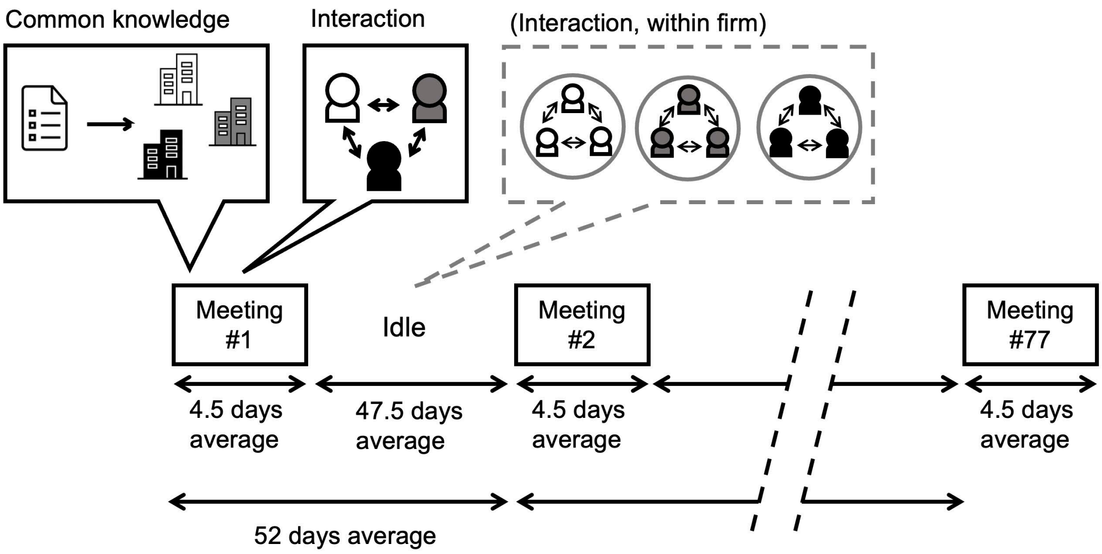
- Interaction takes place via intensive discussions at 3GPP meetings (Kang & Motohashi, 2015), and the publicly available lists of attendants allow us to observe the individuals (and their firms) who engaged in that meeting. For our analysis, we will assume that personal interaction between individuals of firms takes place during these two-monthly meetings (see also Section 8, where we discuss possible limitations of this assumption).

The second aspect that makes 3GPP attractive for our case is that almost all inventions made in this context are patented and can be well tracked to 3GPP work. In the field of mobile telecommunications, there is a race to obtain patents that would be required to implement the standard developed by 3GPP (Bekkers et al., 2002; Leiponen, 2008; Bekkers et al., 2011; Berger et al., 2012). Such patents, known as Standard Essential Patents (SEPs), are well understood to be very attractive, as they can be licensed out, and can secure a strong bargaining position when negotiating cross-licenses to both SEPs and non-SEPs held by other patent owners (Kang & Bekkers, 2015). Tracking patented inventions to 3GPP is possible because all participants are subject to policies that oblige them to disclose whether they believe to own a patent that is or may become essential. These declarations are made publicly available, creating a valuable data source for our study (see Section 5, below, for details). And since patent data also includes inventor names, we can link this data to the list of attendants at 3GPP meetings. Another reason why almost all technical progress is patented in this context is that the process of standards setting implies that companies need to share their knowledge (disclose) in order to get it into the standard, making trade secrets worthless.

Finally, a third aspect is that the training and background of the 3GPP participants that engage in inventions and list as inventors on patents can be regarded as similar (which was one element of our research question). Most of the standardisation engineers have academic backgrounds in electrical engineering, communication engineering, electronic engineering, and information sciences. Students in those fields learn subjects such as information theory, communication theory and systems, signal processing, coding theory and applications, network designs, etc. Therefore, standardisation engineers have a common knowledge pool from academic training.

In the belief that 3GPP, for the reasons listed above, is an appropriate empirical l setting for our research, Figure 1 presents a stylized depiction of the way we test our hypothesis in the 3GPP context.

Figure 1. Exposure and exchange of knowledge in 3GPP



4 Empirical framework

To answer our research question (Section 1) and test the hypothesis developed above, we built a dataset that includes information to create two independent variables related to the two aspects of our research question and includes patent data that allows for observing parallel inventions. Eventually, we test our hypotheses according to the following regression:

$$Y_{ab} = \beta_0 + \beta_1 CEX_{ab} + \beta_2 IIA_{ab} + \alpha T + \delta_1 P_a + \delta_2 P_b + u_{ab} \quad (1)$$

Here, Y_{ab} represents a parallel invention involving the pair of patents *a* and *b* (see Sections 5.2 and 5.3 for details). The first independent variable, CEX_{ab} *(common exposure),* is a dummy variable equaling 1 when the inventors of patents a and b were both exposed to the same, common set of information, which, in line with the previous paragraph, we can observe as both patents being filed after the same 3GPP meeting at which this information was made available to the companies/inventors. The second independent variable, IIA_{ab} *(inventor interaction)*, is a dummy variable equaling 1 when there has been personal interaction between inventors of patents a and b, which, in line with the previous paragraph, we can observe as the inventor(s) of both these patents being physically present at the same 3GPP meeting. It is important to note that we consider this second independent variable as an additional effect to the first – if there is interaction between inventors at a meeting, then there must have been a meeting so the first independent variable must be true as well. The regression also controls for time trends, by including year controls T. Moreover, we include patent-level variables control (P_a and P_b) for patent-specific characteristics otherwise unobserved, such as the patent value proxied as the normalized 7-year forward citations score excluding self-citations.

More details on the outcome variables, independent variables of interest, and controls are provided in the next section.

5 Dataset and main variables

5.1 Dataset construction

Our dataset combines three main types of data sources: meeting data from 3GPP, data on patents declared as potential essential from ETSI, and detailed patent data from the PATSTAT database as well as from additional OECD and USPTO patent datasets. This section provides further information on the creation and properties of our dataset.

Section 3 already introduced 3GPP. In this consortium, most participants are firms, although also other types of entities (e.g., public research labs and government bodies) can participate as long as they are members of one of the collaboration regional standard-setting 3GPP is made of.[1] Attracting thousands of engineers working on different topics within mobile communications, 3GPP formed Technical Specification Groups (TSG), which each may be composed of multiple Working Groups (WG), in which the technical work is carried out. In this paper, we focus data collection on TSG RAN WG1, an important group where fundamental radio technologies on the so-called physical layer are defined. This working group organises frequent meetings – almost bi-monthly. Our 3GPP data builds upon the data collection presented in [reference removed for peer reviewing]. It covers all 77 meetings (on average 4.5 days long), held between its establishment in January 1999 up to February 2010, at quite regular spacing in time of approximately 2 months, and we collected and cleaned information of 939 individual participants at these meetings, affiliated with 53 different firms.

As explained in Section 3, all 3GPP participants are subject to policies that oblige them to disclose whether they believe to own a patent that is or may become essential. This obligation is not within 3GPP itself but exists within the partnering organisations via which they need to participate. The rationale and specifics of such policies have been the topic of many studies; for further details we refer to Contreras, 2013 and Ohana & Biddle, 2018. While there are several partnering organisations, Europe-based ETSI has by far the largest and most detailed database of such

[1] For convenience, this paper simply refers to 'firms'.

disclosures (Baron & Pohlmann, 2018; Bekkers et al., 2020; Jones et al., 2021). Considering that across all the disclosed patent families, the US is the jurisdiction with the highest coverage, we extracted from this database and cleaned up all related 14,510 patent applications applied at the USPTO,[2] covering the time period of our 77 3GPP meetings

For each of these 14,510 disclosed standard essential patents, detailed patent data (patent families, inventors assignees, citations, etc.) was obtained from the EPO PATSTAT patent database (Autumn 2020 edition). We complemented this data with patent quality indicator data by the OECD (Squicciarini et al., 2013), as well as USPTO Patent Prosecution Research Data set (Lu et. Al., 2017) (for more details on the latter dataset, see at Section 5.2). Also, for the 939 individual participants at 3GPP and the 53 different firms they are associated with, we identified their USPTO patent applications in PATSTAT.

5.2 Outcome variable: similar inventions

Most common approaches to identifying similar inventions are based on text analysis. Conventional models measure text similarity are based on keyword frequency and co-occurrence measures (Arts et al, 2018, 2021), but lack the ability to understand contexts and meanings. Recently, AI models were developed to address this shortcoming, such as the Bidirectional Encoder Representations from Transformers (BERT) developed by Google (Devlin et al., 2019). Such systems deal with context and meaning by using multidimensional vectors, a famous example being "King – Man + Woman = Queen" (Baldwin, 2015). Reimers and Gurevych (2019) applied the bi-encoder design to sentence similarity tasks and developed a Sentence-BERT (SBERT) model that lessens heavy calculations while maintaining the accuracy of BERT.

For our study, we employed an improved Sentence-BERT model specialized for patent documents, developed by researchers from Aalborg University called PatentSBERTa (Bekamiri et al, 2021). This model takes into account the specific language and structure used in patent documents. Using PatentSBERTa, we calculated the similarity of patents using the patent claim text that is available in the USPTO Pre-Grant Data set in USPTO's PatentsView[3]. Since this claim data is only available for

[2] They consist of 11,260 patent applications applied for directly at the USPTO and 3,250 US patent applications that were not applied for directly at the USPTO but via the WO/PCT route. The 3,250 WO/PCT-route US patent applications don't match directly with the USPTO Patent Prosecution Research Data, patent claim data, or the OECD Patent Quality database. But, they can be linked via their patent family member in the USPTO with the USPTO Patent Prosecution Research Data, patent claim data, and the OECD Patent Quality database.
[3] https://patentsview.org/download/pg-data-download-dictionary

patents filed in 2005 or later, our dataset for this analysis covers 11,787 US patent applications.[4] The 11,787 patent applications result in a total of 69.5 million patent pairs.[5]

5.3 Outcome variable: identical inventions

For our second outcome variable, we go one step further than the mere similarity between inventions, and want to determine the occurrence of identical inventions. To do so, we exploit the fact that patent examiners, when preparing search reports for new patent applications, specifically distinguish earlier patents that challenge the novelty or inventive step of the patent under examination (Czarnitzki et al, 2011). The presence of such blocking citation means that the two patents in question have a high degree of similarity (and, in fact, means that the patent under examination cannot be granted as applied for, although the application could remove subject matter to obtain a patent with a smaller scope). At the USPTO, a patent (or claim thereof) is rejected on the basis of a lack of novelty under 35 U.S.C. § 102; this is similar to what is called an "X" category reference in WO/PCT international search report, or at the EPO. Furthermore, at the USPTO, a patent is rejected on obviousness (i.e., lack of inventive step) under 35 U.S.C. § 103; this is similar to what is called a "Y" category reference in WO/PCT international search report or at the EPO (USPTO, 2019: p. L-19; EPO, 2020: §9.2.1). For any given patent, information on whether a particular citation challenged novelty and/or inventive step of an applied patent can be found in the related Office Action documents in the public PAIR database of the USPTO (see USPTO, 2010). Until recently, it would take a lot of manual work to collect such unstructured, in-text information for larger sets of patents. In 2017, staff of the USPTO Office of the Chief Economist created a novel dataset on office reject actions and their causes, called the USPTO Patent Prosecution Research Data (PPRD) dataset (Lu et. Al., 2017), and in our study, we use this dataset to observe blocking citations[6]. This dataset is generated using machine-learning techniques and covers all relevant mailed office actions in the period between 2008 and mid-2017. Note that this effectively restricts our analysis to the time period from 2008 (start availability of PPRD data) to February 2010 (the last 3GPP TSG RAN WG1 part of our dataset). While this is a short period, the large number of patents in our dataset and the high frequency of 3GPP meetings leave us with ample data to carry out our analysis.

---

[4] Note that our 11,787 US patent applications include continuing applications data which are filed in or after 2005, have a priority dates ae before 2005. Continuation patents are often used by firms participating in the standardization as a strategy to obtain patents that are essential to technical standards (Omachi, 2004). Our case is not an exception.
[5] The total numbers of pairs for n data points equals $n\times(n-1)/2$ pairs, so for our 11,787 patents we have approx. 69.5 million patent pairs.
[6] Unlike for the EPO or the PCT, information on blocking citations for USPTO patents is not available in the PATSTAT database, unless these patents came from the PCT route and have an international search report.

Since this analysis is not dependent on the availability of USPTO Pre-Grant Data set, we can include a longer time span and include 14,510 US patent applications in our dataset. Using the above method, we identified 5347 applications with blocking relationships between them. These pairs form our outcome variable in the analysis in Section 7.

5.4 Patent-level controls

Our analyses include a number of patent controls that be confounding sources of variation. Firstly, we control for time trends, by including the meeting number preceding the oldest patent of the pair (there are approx. 6 meetings per year, so this is more granular than years). Second, the control for the time that lapsed until the second patent in the pair, again measured by the meeting numbers. Moreover, we include patent-level variables control for patent-specific characteristics otherwise unobserved, such as the patent value proxied as the normalized 7 year forward citations score, the size of the patent family, and the number of claims. For each observation, we include these control variables for both patents of the pair (P_x and P_y). The patent-level variables are all obtained from the OECD Patent Quality dataset (Squicciarini et al., 2013).

6 Main analysis: Similar inventions

In this section, we present our results investigating similar inventions using the AI-based similarity measure as introduced in Section 5.2, above. Turning now to our analysis outcomes, we first look at the descriptive characteristics of our dataset (Table 1) and correlations between all variables (Table 2). Unsurprisingly, the is quite a lot of correlation between the control variables concerning patent quality; after all, the different measured dimensions of patent quality are often strongly related (Gambardella et al, 2008).

Table 1: Descriptive statistics of the dataset for the 'similar inventions' analysis

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
PatentSBerta	63,598,603	0.605655	0.085727	0.075	1
Common _exposure (CEX)	63,598,603	0.023197	0.15053	0	1
Inventor _interaction (IIA)	63,598,603	0.002542	0.050352	0	1
Meeting #	63,598,603	40.97815	9.264258	0	59.5
Meeting_gap (conf2 - conf1)	63,598,603	6.569517	12.0161	-57	59.5
Patent a fwd_cits7	63,598,603	23.69602	46.90869	0	1283
Patent b fwd_cits7	63,598,603	24.47055	46.81946	0	1283
Patent a family_size	63,598,603	5.359195	5.059932	1	32
Patent b family_size	63,598,603	5.338993	5.185586	1	32
Patent a claims	63,598,603	12.61906	15.14831	1	231
Patent b claims	63,598,603	12.35937	13.7409	1	231

Table 2: Correlation table for the 'similar inventions' analysis

Variable	Correlation									
PatentSBerta	1									
Common _exposure (CEX)	0.0047	1								
Inventor _interaction (IIA)	0.0136	0.3276	1							
Meeting #	0.0776	0.0692	0.0391	1						
Meeting_gap (conf2 - conf1)	-0.019	-0.0843	-0.0276	-0.6796	1					
Patent a fwd_cits7	0.0298	-0.0111	-0.0005	-0.1815	0.1325	1				
Patent b fwd_cits7	0.0097	-0.0042	0.0014	0.0137	-0.1415	-0.0014	1			
Patent a family_size	0.0193	-0.0146	-0.0073	-0.2015	0.1416	0.3581	-0.0022	1		
Patent b family_size	-0.0014	0.0069	0.0027	0.014	-0.1408	-0.0024	0.3228	-0.009	1	
Patent a claims	0.0554	0.0026	-0.0006	-0.0359	0.005	0.2445	-0.0018	0.376	-0.0024	1

6.1 *Results for similar inventions*

As the outcome variable in this analysis, the PatentSBERTa similarity score, is a continuous variable, we employ a fractional logit regression specification. The results are shown Table 3, which shows marginal effects.

Column (1) is the regression model with only the first independent variable (common exposure, CEX) present. The marginal effect is positive and significant: common exposure increases the likelihood of a parallel invention by 0.03 percentage points when the independent variable changes from zero to one. Column (2) adds the time dimension controls, reduced the magnitude but still shows a positive and significant effect of out independent variable. When we add all the patent quality control variables (Column 3) the findings remain stable. Hence, we can accept Hypothesis 1: If firms are exposed at a given meeting to the same set of knowledge about challenges and goals, the likelihood increases that after that meeting, these firms bring forward similar inventions.

Columns (4-6) in Table 3 present the result for the second independent variable (inventor interaction, IIA). Also here, the marginal effect is positive and significant, regardless of adding controls in the same fashion as above. The marginal effect we observe now is much stronger as for common exposure, from 1.8 to 2.3 percentage points (so, the difference is 9 to 18 times). Hence, we can accept also Hypothesis 2: If individuals of different firms are personally participating at those given meetings, the likelihood increases that after that meeting, these individuals bring forward similar inventions. Note that we do not perform an analysis with both CEX and IIA because the latter already implies the first (i.e., if there is inventor interaction at meetings, then there also has been common exposure at these meetings).

Table 3: Fractional logit regressions (marginal effects). Dependent variable: PatentSBERTa patent similarity score

	(1)	(2)	(3)	(4)	(5)	(6)
Common _exposure (CEX)	0.0027 [37.75]***	0.001 [13.64]***	0.001 [14.17]***			
Inventor _interaction (IIA)				0.0234 [117.08]***	0.0184 [91.78]***	0.0178 [88.92]***
Meeting #		0.0011 [704.39]***	0.0012 [721.88]***		0.0011 [702.13]***	0.0012 [719.64]***
Meeting_gap		0.0004 [370.22]***	0.0004 [344.23]***		0.0004 [370.10]***	0.0004 [344.03]***
Patent a: fwd_cits7			0.0001 [223.74]***			0.0001 [223.26]***
Patent a: family_size			0.0000 [43.43]***			0.0001 [43.53]***
Patent a: claims			0.0003 [378.03]***			0.0003 [378.14]***
Patent b: fwd_cits7			0.0000 [52.68]***			0.0000 [52.50]***
Patent b: family_size			-0.0003 [-129.27]***			-0.0003 [-129.37]***
Patent b: claims			0.0004 [528.00]***			0.0004 [527.89]***
N	63,598,603	63,598,603	63,598,603	63,598,603	63,598,603	63,598,603

7 Main analysis: Identical inventions

In this section, we present our results investigating identical inventions using the patent examiner X-Y blocking citations (introduced in Section 5.3 above) as the outcome variable. In terms of independent variables and controls, the analysis is similar to that presented in Section 6, above. As this dataset has a different number of total observations (see above), we again present the descriptive characteristics of our dataset (Table 4) and correlations between all variables (Table 5). In these tables, 'Blocking_cit' represents the outcome variable presented in this section, whereas 'Any-cit' is the outcome variable used for the robustness test presented in Annex A.

Table 4: Descriptive statistics for the 'identical inventions' analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
Any-cit	12,902,460	0.000329	0.018133	0	1
Blocking_cit	12,902,460	0.0001	0.010022	0	1
Common _exposure (CEX)	12,902,460	0.048969	0.215803	0	1
Inventor _interaction (IIA)	12,902,460	0.007758	0.087737	0	1
Meeting #	12,902,460	51.2706	3.071257	47	59.5
Meeting_gap (conf2 - conf1)	12,902,460	2.241004	4.290131	-12.5	12.5
Patent a					
- fwd_cits7	12,902,460	18.05416	27.15476	0	356
- family_size	12,902,460	4.250456	5.450398	0	27
- claims	12,902,460	11.45191	15.18933	0	141
Patent b					
- fwd_cits7	12,902,460	20.66979	29.05446	0	356
- family_size	12,902,460	4.726902	4.983562	0	27
- claims	12,902,460	12.41338	13.69101	0	141

Table 5: Correlation table for the ‘identical inventions’ analysis

Variable	Correlation											
Any-cit	1											
Blocking_cit	0.5525	1										
Common_exposure (CEX)	0.0022	0.0015	1									
Inventor_interaction (IIA)	0.0054	0.003	0.3897	1								
Meeting #	-0.0027	-0.0007	0	0.0187	1							
Meeting_gap (conf2 - conf1)	-0.0002	0.0002	0	-0.0462	-0.6314	1						
Patent a fwd_cits7	0.025	0.0144	0	0.0085	-0.0658	0.0668	1					
Patent b fwd_cits7	0.0122	0.006	0	0.0171	0.0179	-0.0403	-0.0032	1				
Patent a family_size	0.0116	0.0055	-0.0231	0	-0.1694	0.1459	0.4305	-0.0042	1			
Patent b family_size	0.0116	0.0037	0.0392	0	0.0113	-0.2249	-0.0062	0.3327	-0.013	1		
Patent a claims	0.0111	0.0061	0	-0.0049	-0.0763	0.0542	0.3752	0.003	0.4293	0.0029	1	
Patent b claims	0.0056	0.0039	0	0.0077	0.0105	0.0121	0.005	0.2681	0.0034	0.2792	-0.0082	1

Turning now to our main analysis, shown in Table 6, where Column (1) is the regression model with only the first independent variable (common exposure, CEX) present. The marginal effect is positive and significant: common exposure increases the likelihood of a parallel invention by 0.01 percentage points when the independent variable changes from zero to one. Column (2) adds the time dimension controls, which do not impact the magnitude or significance. Adding all the patent quality variables (Column 3), the findings remain stable. Also now, we can accept Hypothesis 1. Columns (4-6) in Table 6 present the result for the second independent variable (inventor interaction, IIA). Also here, the marginal effect is positive and significant, regardless of adding controls in the same fashion as above. The marginal effect we observe is twice as strong as for common exposure. Hence, we can again accept also Hypothesis 2.

Table 6 Logistic regressions (marginal effects). Dependent variable: blocking citation between patent pair a and b

	(1)	(2)	(3)	(4)	(5)	(6)
Common _exposure (CEX)	0.0001	0.0001	0.0001			
	[5.11]***	[5.04]***	[5.10]***			
Inventor _interaction (IIA)				0.0002	0.0002	0.0001
				[9.46]***	[9.44]***	[9.46]***
Meeting #		0.0000	-0.0000		-0.0000	-0.0000
		[-2.14]**	[-2.07]**		[-2.33]**	[-2.27]**
Meeting_gap		-0.0000	-0.0000		-0.0000	-0.0000
		[-0.27]	[-0.03]		[-0.49]	[-0.29]
Patent a: fwd_cits7			0.0000			0.0000
			[26.25]***			[26.12]***
Patent a: family_size			0.0000			0.0000
			[4.72]***			[4.79]***
Patent a: claims			0.0000			0.0000
			[9.67]***			[9.61]***
Patent b: fwd_cits7			0.0000			0.0000
			[14.11]***			[14.01]***
Patent b: family_size			0.0000			0.0000
			[6.57]***			[6.62]***
Patent b: claims			0.0000			0.0000
			[8.43]***			[8.41]***
N	12,902,460	12,902,460	12,902,460	12,902,460	12,902,460	12,902,460

A separate robustness effects analysis, presented in Annex A, confirms our findings. In this robustness test, we observe the occurrence of parallel invention by observing any citation relationship between the two patents in the pair (instead of only blocking patents). We find a similar magnitude of effects for common exposure (CEX), and even stronger effects inventor interaction (IIA) shows the likelihood a bit larger or similar than CEX.

8 Discussion and conclusion

This study aims to explain how the occurrence of parallel invention can be understood by the presence of common knowledge across prospective inventors (like knowledge of the task to be achieved) and by personal interaction taking place between prospective inventors. For the first time, we use an empirical, quantitative approach to study this topic. We do so by using data of 3GPP RAN1 standardisation activities, where engineers compete to create new knowledge, and where shared knowledge and personal knowledge interactions can be observed. We identify parallel inventions in two ways: (1) similar inventions, where our measurements are based on an AI-based approach called PatentSBERTa, and (2) identical inventions, where our measurements are based on the outcome of patent examinations, which contains information on patents that are rejected because exactly that same idea was already contained in another patent. Since the latter is a more stringent measure of parallel inventions, we would expect less strong effects.

We find that:

(1) A pair of patent applications have a higher likelihood of representing parallel inventions of they were both filed after a standardisation meeting where information was shared on the tasks to be achieved. This finding implies that when independent individuals are given with the same information input, they are more likely to create parallel inventions. For all the analyses we performed, this effect is robust: for similar inventions, we estimate this effect to be 0.1 to 0.3 percentage points; for identical inventions, the effect is smaller (as to be expected) at 0.01 percentage points.

(2) On top of the above, a pair of patent applications has a higher likelihood of representing parallel inventions if the inventors of both patents were personally present at the meeting in question. This finding implies that when independent individuals are present in the same space and have the opportunity to communicate, they are likely to create parallel inventions. For all the analyses we performed, this effect is robust: for similar inventions, we estimate this effect to be 1.8 to 2.3 percentage points; for identical inventions, the effect is smaller (as to be expected) at 0.02 percentage points. Overall, these effects are much stronger than those for information sharing, stressing the importance of personal communication.

In short, we see a mechanism in which identical knowledge inputs result in similar new knowledge outputs when independent knowledge creators are educated in similar fields and are given an identical knowledge input. We argue that this mechanism is an unrevealed mechanism of multiple discoveries and inventions.

In terms of academic contribution, our work provides empirical the theoretical work of Amabile (1983, 1988) on knowledge creation at the individual level and that of Nonaka (1994, 2006) on knowledge creation at the organizational level, but also echoes Schumpeter's view of innovation as "new combinations of new or existing knowledge, resources, equipment and so on" (Schumpeter, 1934). In terms of managerial implication, our work demonstrates how exposure to new knowledge and personal contact between inventors of different organisations can be used to increase knowledge production, but also how it can lead to parallel invention (which might be undesired and where a given company at least will want to be the first to apply for patent protection, if applicable).

Being the first empirical study in this field, our study also has limitations. Firstly, our observations are based on 3GPP's standardization, where all inventors have engineering backgrounds, which is characterized by its science base. To understand more about the degree of generatability of our findings, the study would be replicated in other academic fields, such as humanities and social sciences. However, there are some challenges in doing so, especially in funding a specific context that allows to observe phenomena like shared information and communication. Knowledge creation in engineering is based purely on science. This is somewhat different from other academic fields, such as humanities and social sciences, where research may be biased by political correctness or ideology to some extent. Secondly, our approach can't see observe the possibility of mere theft of ideas (see Bekkers, Martinelli, and Tamagni, 2020). While such theft may exist, its occurrence is arguable much lower than the phenomenon we are investigating here – also because collaboration in 3GPP is a multistage process, continuing over decades now, and those willing to engage in idea theft will need to consider the longer-term repercussions. Lastly, in the 3GPP standardisation context in which we carried our study, we cannot observe whether in-between meetings, engineers of different companies came together and discussed their ideas. While competitors coming together in such a context where their ideas (and patents) may compete may seem less likely, yet in 3GPP there may be such collaborations when they prepare joint submissions for an upcoming meeting. In any case, we consider our measured effects as a lower bound of the studied effects, which may be higher if we were able to observe such in-between meetings.

Annex A: Robustness check

In this robustness test, we perform the same regression in Table 6, but now the outcome variable is *any* citation between the patents in a pair (instead of a blocking citation). This is arguable a less strict definition of similarity; after all, such citations could also indicate relevance (for assessing whether a patent meets the patentability criteria) rather than that the cited patent is factually the same invention. Results (marginal effects) are shown in Table 7. Columns (1) to (3) show that for the first variable of interest, common exposure (CEX), we observe the same marginal effects as in the main analysis (as presented in Table 6, above). Columns (4) to (6) show that for the second variable of interest, inventor interaction (IIA), the marginal effects are also positive and significant, as in the main analysis, but have a higher magnitude: the marginal effect for IIA is four to five times higher than that for CEX.

Table 7 Logistic regressions (marginal effects). Dependent variable is any citation between patent pair x and y

	(1)	(2)	(3)	(4)	(6)	(5)
Common _exposure (CEX)	0.0001 [7.77]***	0.0001 [6.44]***	0.0001 [6.84]***			
Inventor _interaction (IIA)				0.0005 [16.97]***	0.0005 [16.39]***	0.0004 [15.07]***
Meeting #		-0.0000 [-11.94]***	-0.0000 [-8.19]***		-0.0000 [-12.12]***	-0.0000 [-8.29]***
Meeting_gap		-0.0000 [-7.29]***	-0.0000 [-2.50]**		-0.0000 [-7.38]***	-0.0000 [-2.52]**
Patent a: fwd_cits7			0.0000 [46.29]***			0.0000 [46.06]***
Patent a: family_size			0.0000 [14.60]***			0.0000 [14.73]***

Patent a: claims			0.0000 [15.22]***			0.0000 [15.10]***
Patent b: fwd_cits7			0.0000 [27.09]***			0.0000 [26.88]***
Patent b: family_size			0.0000 [25.73]***			0.0000 [25.93]***
Patent b: claims			0.0000 [8.65]***			0.0000 [8.52]***
N	12,902,460	12,902,460	12,902,460	12,902,460	12,902,460	12,902,460

References

Abdelkafi, N., Bekkers, R.N.A., Bolla, R., Rodriguez-Ascaso, A. & Wetterwald, M. (2021) Understanding ICT Standardization: Principles and Practice (2nd edition). Sofia Antipolis, France: ETSI

Amabile, T.M. (1983) The social psychology of creativity: A componential conceptualization, Journal of Personality and Social Psychology, 45(2), 357–376.

Amabile, T.M. (1988) A model of creativity and innovation in organizations, Research in Organizational Behavior 10, 123-167.

Arts, S., Cassiman, B., & Gomez J.C. (2018) Text matching to measure patent similarity, Strategic Management Journal 39(1), 62–84.

Arts, S., Hou, J., & Gomez J.C. (2021) Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures, Research Policy 50(2), 104144.

Baker, B.H. (2000) The gray matter: the forgotten story of the telephone. St. Joseph (MI): Telepress.

Baldwin, T. (2015). King – Man + Woman = Queen: The Marvelous Mathematics of Computational Linguistics. MIT Technology Review, September 17, 2015

Bardi, J.S. (2009) The Calculus Wars: Newton, Leibniz, and the Greatest Mathematical Clash of All Time. New York: Basic Books.

Baron, J. & Gupta, K. (2018) Unpacking 3GPP standards, Journal of Economics & Management Strategy 27(3), 433-461.

Baron, J. & Pohlmann, T. (2018) Mapping standards to patents using declarations of standard-essential patents, Journal of Economics & Management Strategy 27(3), 504–534.

Baye, M.R., & Hoppe, H.C. (2003) The strategic equivalence of ren-seeking, innovation, and patent-race games, Games and Economics 44(2), 217–226.

Bekamiri, H., Hain, D.S., & Jurowetzki, R. (2021) PatentSBERTa: A Deep NLP based Hybrid Model for Patent Distance and Classification using Augmented SBERT. arXiv preprint arXiv:2103.11933.

Bekkers, R., Martinelli, A., Tamagni, F. (2020). The impact of including standards-related documentation in patent prior art: Evidence from an EPO policy change. Research Policy 49(7), 1040071. https://doi.org/1016/j.respol.2020.104007

Bekkers, R., Bongard, R., & Nuvolari, A. (2011) An empirical study on the determinants of essential patent claims in compatibility standards, Research Policy 40(7), 1001–1015.

Bekkers, R., Duysters, G. & Verspagen, B. (2002) Intellectual property rights, strategic technology agreements and market structure: The case of GSM, Research Policy 31(7), 1141-1161

Bekkers, R.N.A. (2001). Mobile Telecommunications Standards: GSM, UMTS, TETRA and ERMES. Boston, MA: Artech House. ISBN: 978158053250

Berger, F., Blind, K., & Thumm, N. (2012) Filing behaviour regarding essential patents in industry standards, Research Policy 41(1), 216–225.

Bikard, M (2020) Idea twins: Simultaneous discoveries as a research tool, Strategic Management Journal 41(8), 1528–1543. https://doi.org/10.1002/smj.3162

Brooks, J. (1975). Telephone: The first hundred years. New York: Harper & Row.

Brown, B.L. (2020) The Bell versus Gray telephone dispute: resolving a 144-year-old controversy, Proceedings of IEEE 108(11), 2083–2096. 10.1109/JPROC.2020.3017876

Carpenter, M., Narin, F., and Woolf, P. (1981) Citation rates to technologically important patents, World Patent Information 3(4), 160–163. https://doi.org/10.1016/0172-2190(81)90098-3.

Contreras, J.L. (2013) Technical standards and ex ante disclosure: Results and analysis of empirical study, Jurimetrics 53(2), 163–211

Czarnitzki, Dirk and Czarnitzki, Dirk and Hussinger, Katrin and Leten, Bart, The Market Value of Blocking Patent Citations (March 5, 2011). ZEW - Centre for European Economic Research Discussion Paper No. 11-021, Available at SSRN: https://ssrn.com/abstract=1803315 or http://dx.doi.org/10.2139/ssrn.1803315

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186, Minneapolis, Minnesota, June 2019

EPO (2020). Particularly relevant documents (Section 9.2.1). In: EPO Guidelines for Examination. Retrieved 6 June 2022 from https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_x_9_2_1.htm

Evenson, A.E. (2000) The Telephone Patent Conspiracy of 1876: The Elisha Gray-Alexander Bell Controversy and Its Many Players. Jefferson, NA: McFarland & Company

Evenson, A.E. (2001) The Telephone Patent Conspiracy of 1876: The Elisha Gray-Alexander Bell Controversy and Its Many Players. Jefferson, North Carolina: McFarland Publishing.

Farr, J.L., & Ford, C.M. (1990) Individual innovation, In West, M.A., & Farr, J.L. (Eds.), Innovation and Creativity at Work, Psychological and Organizational Strategies (pp. 63-80). New York, NY: John Wiley & Sons.

Foley, V., Palmer, G., & Soedel, W. (1985) The Crossbow. Scientific American, 252(1), 104-111.

Fudenberg, D., Gilbert, R., Stiglitz, J., & & Tirole, J. (1983) Preemption, leapfrogging and competition in patent races, European Economic Review 22(1), 3–31.

Gambardella, A., Harhoff, D., & Verspagen, B. (2008) The value of European patents. European Management Review 5 (2), 69–84.

Groenewoudt, A.C., Rooks, G., & van Gool, P.J.R. (2019) When problems lead to ideas: The roles of daily vigor and social interactions, Journal of Creative Behavior 53(3), 286-297.

GSMA (2022). The Mobile Economy 2022. Retrieved 20 May 2022 from https://www.gsma.com/mobileeconomy/wp-content/uploads/2022/02/280222-The-Mobile-Economy-2022.pdf

Gurteen, D. (1998) Knowledge, creativity and innovation, Journal of Knowledge Management 2(1), 5-13.

Hammond, M.M., Neff, N.L., Farr, J.L., Schwall, A.R., & Zhao, X. (2011) Predictors of individual-level innovation at work: A meta-analysis, Psychology of Aesthetics, Creativity, and the Arts 5(1), 90–105.

Heap, J.P. (1990) The Management of Innovation & Design. London: Cengage Learning.

Jehn, K.A., Northcraft, G.B., & Neale, M.A. (1999) Why differences make a difference: A field study of diversity, conflict and performance in workgroups, Administrative Science Quarterly 44(4), 741-763.

Johansson, M., Kärreman, M., & Foukaki, A. (2019) Research and development resources, coopetitive performance and cooperation: The case of standardization in 3GPP, 2004–2013, Technovation 88, 102074.

Jones, S.L., Leiponen, A., & Vasudeva, G. (2021) The evolution of cooperation in the face of conflict: Evidence from the innovation ecosystem for mobile telecom standards development, Strategic Management Journal 42(4), 710–740.

Kang, B., & Bekkers, R. (2015) Just-in-time patents and the development of standards, Research Policy 44(10), 1948–1961.

Kang, B., & Motohashi, K. (2015) Essential intellectual property rights and inventors' involvement in standardization, Research Policy 44(2), 483–492.

Karki, M.M.S. (1997) Patent citation analysis: A policy analysis tool, World Patent Information 19(4), 269–272.

Khamara, E.J. (2006) Space, Time and Theology in the Leibniz-Newton Controversy. Berlin: De Gruyter.

Kim, E. (2017) Workshop design for enhancing the appropriateness of idea generation using analogical thinking, International Journal of Innovation Studies 1(2), 134-143.

Kuhn, T.S. (1959) Energy conservation as an example of simultaneous discovery, In Clagett, M. (Eds.), Critical Problems in the History of Science: Proceedings of the Institute for the History of Science (pp. 321-356). Madison, WI: The University of Wisconsin Press.

Lamb, D., & Easton, S.M. (1984) Multiple Discovery: The Pattern of Scientific Progress. Farnham, England: Avebury Publishing Company.

Leiponen, A.E. (2008) Competing through cooperation: The organization of standard setting in wireless telecommunications, Manage Science 54(11), 1904–1919.

Lu, Qiang; Myers, Amanda; Beliveau, Scott (2017). USPTO Patent Prosecution Research Data: Unlocking Office Action Traits. USPTO Economic Working Paper No. 2017-10. Retrieved 6 June 2022 from https://www.uspto.gov/ip-policy/economic-research/research-datasets/office-action-research-dataset-patents

McAdam, R., & McClelland, J. (2002) Individual and team-based idea generation within innovation management: organizational and research agendas, European Journal of Innovation Management 5(2), 86–97.

Nagaoka, S., Motohashi, K., & Goto, A. (2010) Patent statistics as an innovation indicator, In Hall, B., & Rosenberg, N. (Eds.), Handbook of the Economics of Innovation Vol. 2. (pp. 1083-1127). Amsterdam: Elsevier.

Nonaka, I. (1994) A dynamic theory of organizational knowledge creation, Organization Science 5(1), 14-37.

Nonaka, I., & Takeuchi, H. (1995) The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. New York: Oxford University Press.

Nonaka, I., Von Krogh, G., & Voelpel, S. (2006) Organizational knowledge creation theory: Evolutionary paths and future advances, Organization Studies 27(8), 1179-1208.

Nonaka, I., Konno, N. (1998). The concept of Ba: Building a foundation for knowledge creation. California Management Review, 1998, (3), pp. 40–54

Ogburn, W.F., & Thomas, D. (1922) Are inventions inevitable? A note on social evolution, Political Science Quarterly 37(1), 83–98.

Ohana, G., & Biddle, C.B. (2018) The disclosure of patents and licensing terms in standards development, In Contreras, J.L. (Eds.), The Cambridge Handbook of Technical Standardization Law: Competition, Antitrust, and Patents. (pp. 244-260). Cambridge: Cambridge University Press.

Oldham, G.R., & Cummings, A. (1996) Employee creativity: Personal and contextual factors at work, Academy of Management Journal 39(3), 607-634.

Omachi, M. (2004). Emergence of Essential Patents in Technical Standards: Implications of the Continuation and Divisional Application Systems and the Written Description Requirement. IIR Working Paper WP#05-02. Retrieved from http://hermes-ir.lib.hit-u.ac.jp/rs/bitstream/10086/15993/1/070iirWP05-02.pdf

Randy, A. (2008). July 1, 1858: Darwin and Wallace Shift the Paradigm. Wired, July 1, 2008. Retrieved 10 June 2022 from https://www.wired.com/2008/07/dayintech-0701/

Reimers, N., & Gurevych, I. (2019) Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 3982–3992, Hong Kong, China, Nov. 2019.

Bekkers, R., Raiteri, E., Martinelli, A., & Tur, E.M. (2020) Landscape Study of Potentially Essential Patents Disclosed to ETSI. Nikolaus Thumm (ed.). Publications Office of the European Union, Luxembourg, JRC121411, doi:10.2760/313626, ISBN 978-92-76-20786-3.

Schumpeter, J. (1934) The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle. Cambridge, MA: Harvard University Press.

Shulman, S. (2009) The Telephone Gambit: Chasing Alexander Graham Bell's Secret. New York: Norton & Company.

Simonton, D.K. (1979) Multiple discovery and invention: Zeitgeist, genius, or chance? Journal of Personality and Social Psychology 37(9), 1603–1616.

Simonton, D.K. (1986) Multiples, Poisson distributions, and chance: An analysis of the Brannigan-Wanner model, Scientometrics 9(3-4), 127-137.

Squicciarini, Mariagrazia; Dernis, Hélèn, Criscuolo, Chiara (2013). Measuring Patent Quality: Indicators of Technological and Economic Value. OECD Science, Technology and Industry Working Papers DSTI/DOC(2013)3.

Stokes, T.D. (1986) Reason in the Zeitgeist, History of Science 24(2), 111-123.

Titus, P.A. (2000) Marketing and the creative problem-solving process, Journal of Marketing Education 22(3), 225-235.

Trajtenberg, M. (1990) A penny for your quotes: Patent citations and the value of innovations. RAND Journal of Economics 21(1), 172–187.

USPTO (2010). Public PAIR Quick Start Guide. Retrieved 6 June 2022 from https://www.uspto.gov/sites/default/files/patents/process/status/private_pair/PAIR_Glossary.pdf

USPTO (2019) Claim drafting. Invention-Con 2019 Pre-Conference Session, Alexandria, VA, September 12, 2019

USPTO (2019). Manual of Patent Examining Procedure - Consolidated Patent Laws — March 2021 update. Retrieved 6 June 2022 from https://www.uspto.gov/web/offices/pac/mpep/consolidated_laws.pdf

USPTO (2020). U.S. Patent Statistics Chart, Calendar Years 1963 – 2020, Retrieved 8 June 2022 from https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm

Vygotsky, L.S. (1986) Thought and Language. Cambridge, MA: Massachusetts Institute of Technology Press.

Wang, X., Ren, H., Chen, Y., Liu, Y., Qiao, Y. & Huang, Y. (2019) Measuring patent similarity with SAO semantic analysis, Scientometrics 121(1), 1–23.

World Trade Organization Committee on Technical Barriers To Trade, "Second triennial review of the operation and implementation of the agreement on technical barriers to trade, document G/TBT/9," 2000.

Younge, K.A. & Kuhn, J.M. (2016) Patent-to-Patent Similarity: A Vector Space Model, Available at SSRN: https://ssrn.com/abstract=2709238