Scientometrics (2024) 129:3023–3043
https://doi.org/10.1007/s11192-024-05049-5

"Divergent" cross-domain stretching for technology fusion: validating the knowledge partition search model using patent data

Jie Liu[1]

Received: 27 April 2022 / Accepted: 3 May 2024 / Published online: 27 May 2024

Abstract
Technology fusion refers to the phenomenon in which distinct technology domains overlap. Despite its importance in technology innovation and evolution, few studies have explored the general pattern of the cross-domain search process leading to technology fusion. This paper proposes that the stretching between distinct technology domains could be viewed as searching in a two-dimensional knowledge partition landscape and then empirically validates the model based on a large patent dataset derived from the U.S. Patent and Trade Office (USPTO). The findings show that the general pattern of the search processes leading to technology fusion could be viewed as searching across a broad technology scope to identify limited valuable linking points within existing technology domains, and the search processes are mainly "divergent"; that is, innovative agents gradually extend the search scope to pursue new hybrid technologies. However, the cross-domain search would be more targeted if the two technology domains were closer to each other. In addition, compared to searching across a broader technology scope, digging in certain technology areas is more important for the generation of new high-impact hybrid technologies. This study provides a novel perspective for understanding the new knowledge creation process and technology fusion.

Keywords Technology fusion · Search and recombination · Patent data · Knowledge partition search model

Introduction

Technology fusion addresses the phenomenon of technology overlap, which is regarded as a crucial mechanism underlying the emergence of new technologies (Caviggioli, 2016; Curran & Leker, 2011; Kim & Kim, 2012). Technology fusion has recently attracted an increasing amount of attention in the management literature because, in recent years, the miniaturization, digitalization, and architectural changes of increasingly complex products

✉ Jie Liu
jliu@njust.edu.cn

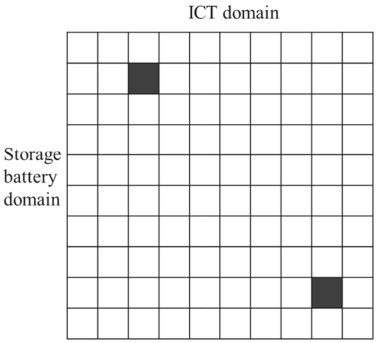
[1] School of Intellectual Property, Nanjing University of Science and Technology, Xiaolingwei Street No. 200, Xuanwu District, Nanjing 210094, Jiangsu, China

have made the overlap of technologies and the recombination of pieces of knowledge belonging to different domains more salient (Ardito et al., 2023; No & Park, 2010). However, it is not easy for innovative agents in one technology domain to search for combination candidates from another complex technology domain; they need to first understand the other domain and then further search that domain (Nakamura et al., 2015). As technology domains become increasingly complex, the cost of cross-domain knowledge collection and integration, as well as the uncertainty of cross-domain searches, becomes more relevant, which may limit innovative agents from exploring new technology. To support innovative agents in such explorations, existing studies have empirically explored the factors that drive technology fusion and have attempted to forecast technology fusion based on different criteria (Kwon et al., 2020; Lee et al., 2021; Sick & Bröring, 2022). However, an empirically validated model of the creation process of technology fusion is still lacking.

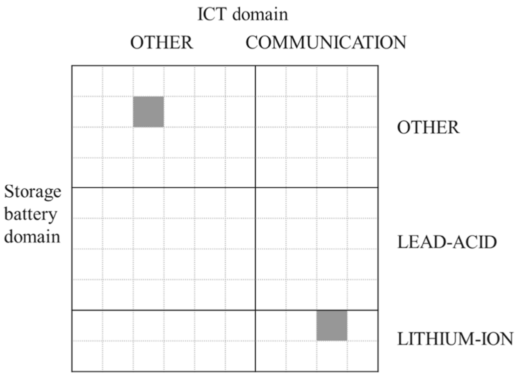
This paper focuses on understanding and monitoring the general dynamics of the knowledge search process of stretching between distinct technology domains. Stretching has been viewed as the driver of technology fusion (Caviggioli, 2016). In this paper, technology fusion refers to the event of a new hybrid technology generation. Previous studies have conceptualized the technology domain as a complex system with a wide range of subdomains (Fleming & Sorenson, 2001; Knudsen & Srikanth, 2014), and the knowledge search process is viewed as containing two main steps, namely, deepening the knowledge on the system to select subdomains and then further searching the target subdomains (Knudsen & Srikanth, 2014; Nakamura et al., 2015). In addition, economists model knowledge as partitions in a state space, which provides a convenient tool for modeling the evolving knowledge of technology systems (Knudsen & Srikanth, 2014; Samuelson, 2004). Borrowing the idea from the conceptualization of the technology system and the knowledge partition model, this paper proposes that the stretching could be viewed as searching in a two-dimensional knowledge partition landscape. Based on the knowledge partition search model, this paper posits that pre-fusion search includes (1) exploring unknown subdomains and (2) digging in known subdomains, which facilitates the recognition and identification of valuable knowledge components across the technology landscape and subsequently enables technology fusion.

To clarify the proposed knowledge partition search model, Fig. 1 provides an example in the context of the example of energy storage (ES) digitalization, i.e., the fusion between information and communication technology (ICT) and ES technology. The process of stretching between distinct technology domains is simplified to searching in a two-dimensional task landscape, which in the ES digitalization example consists of the ICT domain and the storage battery domain. The task landscape is assumed to be a 10×10 matrix in which each cell defines a cross-domain combination candidate, i.e., a potential fusion opportunity with an associated payoff; the valuable fusion opportunities are marked in black in Fig. 1a. At the beginning of the search process, fine-grained partitions of the landscape may not be available, and innovative agents could see only limited choices for each dimension, e.g., the 3×2 choice set (6 subspaces) shown in Fig. 1b, instead of reality (the full 10×10 matrix). In this paper, the pre-fusion search is expected to explore different subspaces and dig in these subspaces to identify valuable candidates.

Given the knowledge partition search model, this paper empirically justifies the conceptual model using patent data, which is based on search and recombination theory. Note that although it may not be explicitly addressed, the search and recombination theory underlies data-driven quantitative technology fusion analysis, in which the technology domain is indicated by the knowledge component set, and technology fusion is identified as the connection between knowledge components in different domains. Prior studies focusing



(a) The task landscape



(b) An example of incomplete knowledge of the task landscape

Fig. 1 The knowledge partition landscape in the context of ES digitalization

on the pattern, the driver, as well as the forecast of technology fusion, mainly address the relevance of knowledge component attributes (Caviggioli, 2016; No & Park, 2010; Sick & Bröring, 2022; Xiao et al., 2022); few studies explore the pattern of the search process leading to technology fusion.

This paper aims to fill this gap in two ways. First, a methodological approach is proposed to combine the knowledge partition search model with data-driven technology fusion analysis based on patent citation and patent co-classification information; this approach is applied to a large patent dataset derived from the U.S. Patent and Trade Office (USPTO). The findings show that the general pattern of the search processes leading to technology fusion could be viewed as searching across a broad technology scope to identify limited valuable linking points within existing technology domains and that the search processes are mainly "divergent"; that is, innovative agents gradually extend the search scope to pursue new hybrid technologies. Second, this paper also investigates the relationship between pre-fusion search process characteristics and the technological impact of the new hybrid technology. Compared to searching across a broader technology scope, digging in certain areas is more important for the generation of new high-impact hybrid technologies, which justifies the proposed conceptual model further and helps guide effective knowledge searches.

The remainder of the paper is organized as follows. Section "Background studies" reviews the related literature. Section "Methodology" describes the methodology. Section "Results" shows the analysis results. Section "Discussion" provides the discussion, and Section "Conclusions" shows the conclusions.

Background studies

Technology fusion

Technology fusion or convergence was first introduced by Rosenberg (1963), who defined it as the process of interdependence between different technologies in the production process. Despite the increasing relevance of fusion to the evolution and radical

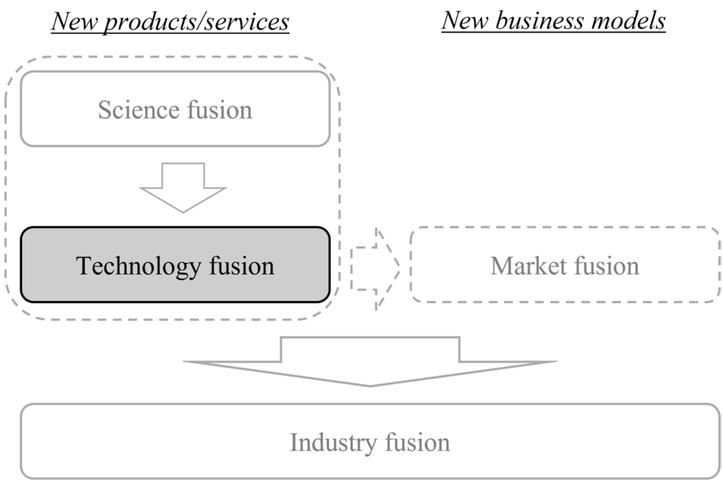


Fig. 2 Science, technology, market, and industry fusion. *Note*: Based on Curran and Leker (2011). Technology fusion has been marked in gray to show that it is the focus of this paper

change of industry, a general definition of the term *fusion* is still lacking; thus, multiple definitions exist (Kangas et al., 2021; Preschitschek et al., 2013). In this paper, the much-cited definition from Curran and Leker (2011) is employed. Thus, fusion is defined as the blurring of boundaries between at least two disjoint areas of science, technology, markets, or industries, which results in the creation of a new (sub) segment as a merger of the parts of the old areas. According to the *loci* of fusion, Curran and Leker (2011) also suggest four stages of fusion, namely, science, technology, market, and industry fusion, as well as the nonlinearity of the fusion process (see Fig. 2).

Science fusion refers to the blurring of different scientific discipline or area boundaries. Technology fusion follows science fusion and may induce market fusion. Industry fusion emerges after science, technologies, and/or market fusion, leading to the merging of companies. Notably, fusion does not necessarily lead to new industries or markets (Curran & Leker, 2011; Kangas et al., 2021). This study focuses on technology fusion, i.e., the blurring of boundaries between disjoint technology areas, which is shown in the phenomenon that two hitherto different industrial sectors come to share common knowledge and technological base (Athreye & Keeble, 2000).

Curran and Leker (2011) note the slight difference between the terms "*convergence*" and "*fusion*". The former refers to a process in which the components in two domains move or stretch to a new and common place; the latter denotes that the components merge in the very same place of at least one of the domains. Similar to Caviggioli (2016), this study does not emphasize this difference and regards these two terms as interchangeable. In addition, scholars have distinguished several additional categories of fusion, such as demand-side fusion, which is associated with the fusion of demand structures, and supply-side fusion, which is related to the sharing of technologies (Bröring & Leker, 2007). This paper focuses on technology-driven supply-side technology fusion.

Search and recombination perspective on technology fusion

Technology evolution has long been viewed as the recombination of existing technologies, and technology innovation is described as the process of searching for valuable combinations of knowledge components (Arthur, 2007; Fleming, 2001; Fleming & Sorenson, 2001; Keijl et al., 2016). Here, a knowledge component denotes a self-standing embodiment of a core concept for a distinct scientific or engineering principle (Galunic & Rodan, 1998; Xiao et al., 2022). The search and recombination perspective helps to unfold an invention's "innards" by putting a spotlight on the knowledge components, including why they are chosen, how they are combined, and the linkage between combination features and value (Xiao et al., 2022). Conceiving that knowledge components can be meaningfully rearranged, search and recombination scholars have also revealed how knowledge residing within an individual, firm, or industry can be extended in new ways and explained the search motivation or behavior of an innovative agent with a particular knowledge base (Xiao et al., 2022).

The search and recombination perspective naturally suits the theme of technology fusion. In fact, the stretching leading to technology fusions could be viewed as the process of searching for combinations of distinct knowledge components belonging to different domains, which paves the way to novel technology domains (Ardito et al., 2023). Indeed, although they may not directly employ the term "fusion" (or "convergence"), many search and recombination scholars conduct innovation studies that are in line with the characterization of technology fusion; i.e., combining two or more distant components may produce a hybrid breakthrough (Keijl et al., 2016; Nemet & Johnson, 2012).

Given the starting point of the knowledge component, search and recombination scholars have explored the characteristics of individuals or sets of knowledge components that may facilitate technology fusion. For example, knowledge components with features such as being nearer to basic research (Jeong & Lee, 2015; Karvonen & Kässi, 2013), having high relatedness (Caviggioli, 2016), and having greater technology search breadth (Ardito et al., 2023) are more likely to be influenced by the technology overlap process.

It is also important to reveal the features of the cross-domain search process leading to technology fusion. For example, Nakamura et al. (2015) propose a knowledge combination model considering knowledge combinations in depth and breadth based on the similarities of two technology domains. Most of the existing search process feature studies address the relationship between features and the performance of combinations. For example, Keijl et al. (2016) find that a combination of components from local, adjacent, and distant knowledge domains has the highest level of technological impact. However, these studies still focus on the features of the consequent combinations, while empirical analysis of the dynamics of the search process leading to the fusion of certain technologies is still lacking.

Search and recombination simulation is a useful method for exploring the relationship between search process dynamics and search efficiency. For example, Knudsen and Srikanth (2014) use an agent-based simulation model to investigate how coordinated exploration by multiple specialists is different from individual searches. Additionally, using agent-based simulation, Puranam and Swamy (2016) explore how the initial representations held by learners influence coupled learning processes; i.e., specialists from different domains learn how to make interdependent choices among alternatives. Moreover, simulation scholars have proposed empirical validation of implications derived from conceptual models.

Measuring technology fusion through patent data

Patent data systematically contain the raw information created by both inventors and expert patent examiners over hundreds of years (Singh et al., 2021), which helps identify general technology changes based on the temporal sequence of patent application and publication, as well as the relation of the patent to other patents and publications (Kim & Kim, 2012). Patents focus on invention activity, which is the mediator of the pursuit of scientific knowledge and product development and is usually the first step toward new technologies (Singh et al., 2021). This feature causes the analysis of patent data to address the problem mainly at the *locus* of technology, which is distinct from the locus of science or industry. Patent data have been found useful in analyzing technology fusion measurements in a number of previous studies (Kim et al., 2014; Kim & Kim, 2012; Klarin et al., 2023). In particular, patents have been employed in the context of technology-driven technology fusion (Caviggioli, 2016). While the most common method for studying technology fusion through patent data is co-classification analysis, patent citation analysis and a combination of co-classification and patent citation have also been used (Kangas et al., 2021).

For co-classification analysis, hierarchical patent classifications, such as International Patent Classification (IPC) codes, are assigned to patent filings according to the technical features of the inventions. The hierarchical structure of IPC codes (or other patent classification systems) facilitates the definition of technology domain and technological distance while also focusing analyses on technology rather than products or markets (Caviggioli, 2016; Nemet & Johnson, 2012). Since the same document can be associated with several classes, co-classification information can be used to identify the relationships between technologies (Karvonen & Kässi, 2013; Kim & Kim, 2012; Park & Yoon, 2014), which includes the overlap of technology domains.

Patent citation networks work as an alternative method for investigating technology fusion (Nemet & Johnson, 2012; No & Park, 2010) since they help to identify technology flows and trajectories (Jaffe & Trajtenberg, 1996; Kim et al., 2014). Technology domains increasingly citing each other's patents are thought to imply overlap between the two technology domains. However, compared to co-classification, patent citation measurement appears more appropriate for, according to Curran and Leker (2011) and Caviggioli (2016), the overlapping process as the stretching of one domain to another rather than to the presence (or emergence) of the technology fusion event. In particular, Caviggioli (2016) has shown that cross-citation between technology domains could work as the driver of technology fusion events.

This paper investigates the dynamics of the stretching process leading to technology fusion based on the combination of patent citation and patent co-classification information. Specifically, this paper introduces a methodology that reveals the dynamics of the stretching process through empirically validating the knowledge partition search model.

Methodology

The two-dimensional knowledge partition search model

The methodology proposed in this paper aims to understand the knowledge search process leading to technology fusion. To do so, a two-dimensional knowledge partition search

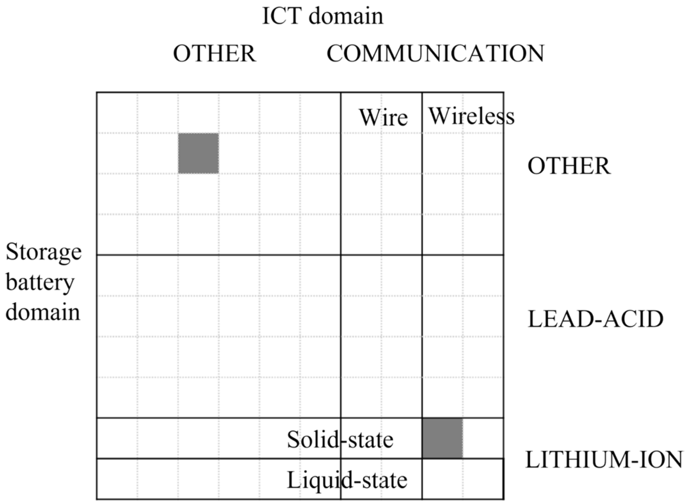


Fig. 3 Dig in a subspace: a new partition structure in the context of ES digitalization

model is employed to analogize the cross-domain search and recombination, i.e., the stretching between two technology domains. The model assumes that what is "known" for agents is represented as a set of partitions of a two-dimensional task landscape. The higher the number of partitions the agents conceive is, the greater the knowledge about the landscape is. In the ES digitalization example, when agents are ignorant about the landscape, only one category for each dimension could be seen; however, if they become more knowledgeable, several categories, such as distinguishing batteries into lead-acid or lithium-ion batteries, could be identified. In the search process, innovative agents may need to refine the partition structure so that it is possible to identify valuable combinations in the task landscape.

The search process proceeds by searching across different subspaces in the current partition structure or by digging in certain subspaces to further partition the space. For example, in the ES digitalization problem, the initial partition structure could be limited to the six subspaces shown in Fig. 1b. The agents could explore different subspaces or select one of them for further investigation, for example, digging in the communication-lithium-ion subspace and discovering not only two different types of lithium-ion batteries, namely, solid- or liquid-state lithium-ion batteries, but also two different types of communication technology, namely, wire- and wireless communication (see Fig. 3).

This paper measures the knowledge search process that matches the abovementioned knowledge partition search model based on patent data and, more importantly, reveals the general pattern of the cross-domain search process leading to technology fusion.

Table 1 Example of the CPC system structures

Hierarchy	Codes	Description
Section	H	Electricity
Class	H01	Electric elements
Subclass	H01C	Resistors
Main group	H01C 10	Adjustable resistors
Subgroup	H01C 10/22	Resistive element dimensions changing gradually in one direction
	H01C 10/50	Structurally combined with switching arrangements

Data

The analysis is based on a dataset derived from the PatentsView platform.[1] The dataset was collected in November 2022 and contains all patents granted by the USPTO since 1976. In this paper, the Cooperative Patent Classification (CPC) system is employed to identify technology domains and technology fusions. The CPC system was developed by the European Patent Office (EPO) and the USPTO to harmonize patent classifications and to replace the former European patent classification system (ECLA) and U.S. patent classification system (USPC). It is divided into nine sections, A-H and Y, which in turn are subdivided into classes, subclasses, main groups, and subgroups (see Table 1) (Cavalheiro et al., 2016; Oh et al., 2020). The CPC system is similar to the IPC system but is more detailed and appears more thoroughly applied, particularly to older patents (Hötte et al., 2021). Only the granted utility patents (for a similar setting, see Singh et al., 2021) applied between 1976 and 2017 are considered in this paper.

Method

Patent data and the knowledge partition search model

Following Caviggioli (2016), the event of technology fusion is identified based on the co-classifications of the CPC; i.e., the first co-occurrence of the CPC code pair is defined as the creation of a new hybrid technology. Specifically, CPC subclasses (4-digit CPCs) are selected among the diverse aggregation levels of analysis to denote technology domains because, similar to IPC codes, the subclass level provides a reasonable and treatable number of technology categories, which also allows sufficient characterization of the technologies (Park & Yoon, 2014).

To explore the search process leading to technology fusion events, in this paper, the cross-citations between the technology domains denoted by the two different CPC subclasses are analyzed in detail. In particular, the hierarchy structure of the CPC system matches the knowledge partition search model well. Given that the CPC subclasses denote technology domains, the CPC main groups (e.g., A01L06) and subgroups (e.g., A11L06/01) are used to indicate the subdomains and components of the technology

[1] https://patentsview.org/download/data-download-tables

domains. Considering the ES digitalization example, the CPC main groups denote the incomplete categories of each domain, such as the subdomain of lithium-ion batteries or communication technology, while the CPC subgroups indicate the components in the (relatively) complete partition structure, such as the solid-state lithium-ion battery in Fig. 3.

Cross-domain search process measurement

Borrowing the related and unrelated variety measurements from Castaldi et al. (2014), in this paper, entropy-based diversification measurements of the cross-citations between two technology domains at two different aggregation levels, i.e., the CPC main group and subgroup levels, are constructed to describe the dynamics of the search process. According to the knowledge partition search model, the main group level diversification measurement is associated with searching across different subspaces of the incomplete partition structure, and the subgroup level diversification denotes the level of digging in the subspace for refining the partition structures.

Specifically, the focal technology fusion samples are the 4-digit CPC pairs that were first combined between 2011 and 2013; that is, the subclass CPC pairs were first co-classified in patents applied between 2011 and 2013. Cross-citations between each pair of CPC subclasses made during 2001–2010 are derived to measure the stretching process. That is, for each pair of CPC subclasses “X” and “Y”, patents applied during the 2001–2010 period in the CPC subclass “X” citing the patents in the CPC subclass “Y” (or “Y” citing “X”) are collected for analysis.

To clarify the process of measuring the cross-citation diversifications, take an example of the CPC subclass pair A01H-A01D, assuming that they were first co-classified in the year 2011. The citation records in which the citing patents in CPC subclass A01H (A01D), the cited patents in CPC subclass A01D (A01H), and the citing patents are applied in the prior two 5-year periods, that is, 2001–2005 and 2006–2010, are collected. The CPC main groups and subgroups in subclass A01H or A01D of the collected citing and cited patents are listed. In this way, two patent-CPC lists are generated for the focal CPC subclass pair and for each period, i.e., the A01H-associated patent-CPC list and the A01D-associated patent-CPC list for the period 2001–2005 and 2006–2010. Figure 4 illustrates the above-mentioned process.

For each list, the main group level diversification, which implies a search across subspaces, is given as follows:

$$SD_l = \sum_{g=1}^{G} P_g ln\left(\frac{1}{P_g}\right) \tag{1}$$

where P_g is the share of patents in CPC main group g in the list. The value of SD_l is bounded from below by zero and has a maximum of $ln(G)$. SD_l is zero if $P_g = 1$ for a single value of g. In the above example, if the A01H-associated patent-CPC list contains only a single CPC main group, then the corresponding SD_l is zero.

One patent may be assigned to more than one CPC main group in the focal CPC subclass. To make the sum of P_g equal to 1, in this paper, the weight of one patent is assigned based on the share of the CPC subgroup. For example, if the CPC subgroups of one patent are A01H1/04, A01H1/05, and A01H4/07, then 0.667 of the patent is assigned to A01H1,

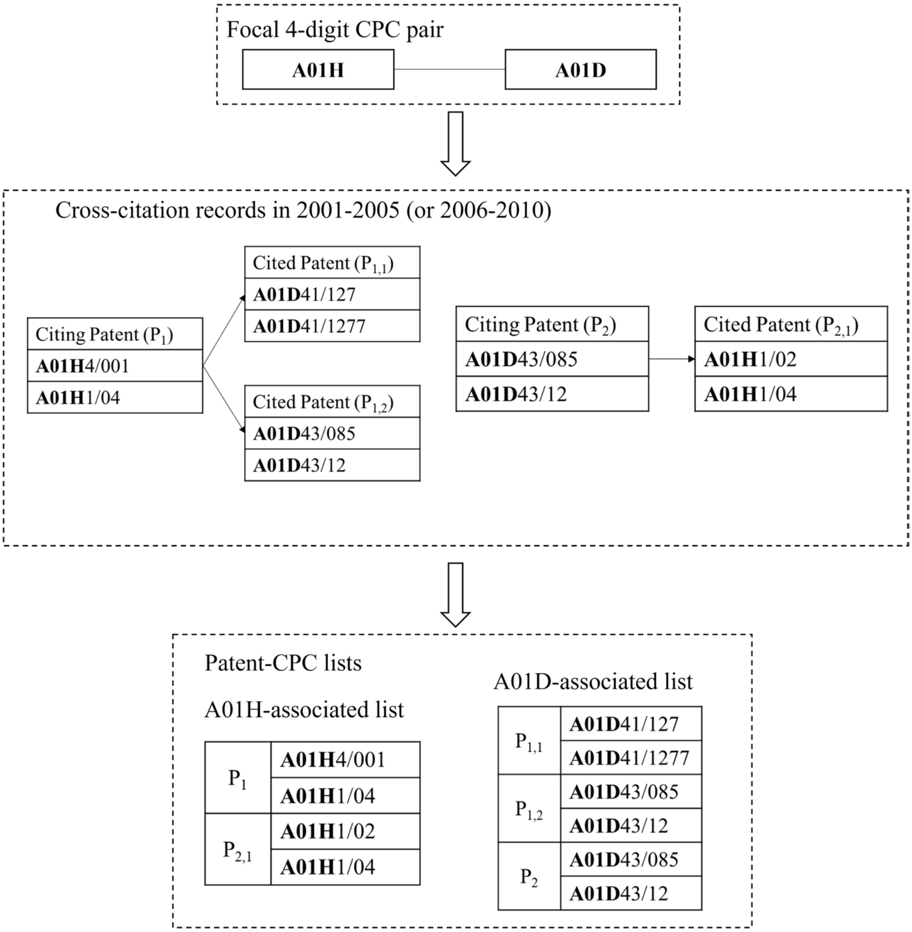


Fig. 4 The process of generating the patent-CPC list

and 0.333 is assigned to A01H4 (also 1/3 for each CPC subgroup). The weighted sum of entropy within each CPC main group is used to measure subgroup level diversification, and following Castaldi et al. (2014), it is given by the following:

$$CD_l = \sum_{i=1}^{I} p_i ln\left(\frac{1}{P_i}\right) - \sum_{g=1}^{G} P_g ln\left(\frac{1}{P_g}\right) \tag{2}$$

where p_i is the share of patents in subgroup i in the list. In fact, Eq. (2) implies that one can consider entropy at the lowest level of aggregation as the sum of entropy within groups at a higher level of aggregation and entropy between these groups. The *SD* or *CD* of a certain 4-digit CPC pair is the sum of the corresponding SD_l or CD_l of two lists.

The *CD* measure reflects the diversification of the search scope across knowledge components within even narrower technology areas, i.e., the CPC main group level

Table 2 Summary results on the sample fusion of USPTO patents

Row	Identified fusions	Number	Percent (%)
(1)	Total fusions made during the 2011–2013 period	6246	100.00
(2)	Fusions with cross-citing during the 2001–2011 period	5808	92.99
(2.1)	Fusions with cross-citing during the 2001–2005 period	5275	84.45
(2.2)	Fusions with cross-citing during the 2006–2010 period	5596	89.59
(2.3)	Fusions with cross-citing during the 2001–2005 and 2006–2010 periods	5063	81.06

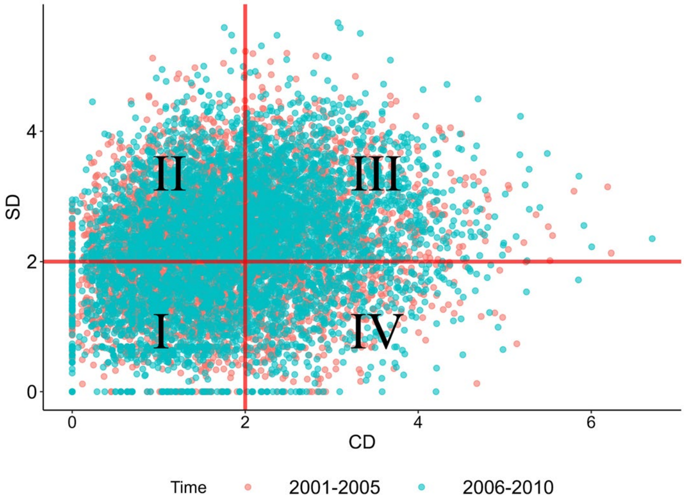


Fig. 5 The distribution of fusions across the levels of SD and CD

technology domains (corresponding to the "subspace" in the knowledge partition landscape), which reflect the level of digging in certain subspaces. In contrast, for *SD*, the entropy at the CPC main group level reflects the extent to which innovative agents search across the different technology subspaces.

Results

Backward-looking on technology fusions

For each pair of CPC subclasses "X" and "Y", which were first co-classified in the sampled years (2011, 2012, and 2013),[2] the cross-citation records between the two CPC subclasses made from 2001 and 2010 are identified; that is, the citing-cited patent relations in which citing patents are in CPC "X" and cited patents are in CPC "Y", or citing patents are in

[2] CPCs in the "Y" section are not considered in this study. Y sections do not indicate separate technological classes but are additional tags attached to patents by examiners to tag some special technical subjects.

Table 3 The number of fusions across four categories in two periods

Category	Number of fusions (2001–2005)	Number of fusions (2006–2010)
I	2148	1862
II	1459	1528
III	1088	1523
IV	580	683
Total	5275	5596

Table 4 The number of category changes in fusions from the 2001–2005 period to the 2006–2010 period

The category in the 2001–2005 period	The category in the 2006–2010 period				
	I	II	III	IV	Total
I	1120	466	158	212	1956
II	221	789	381	55	1446
III	23	144	839	82	1088
IV	102	42	135	294	573
Total	1466	1441	1513	643	5063

CPC "Y" and cited patents are in CPC "X". As mentioned above, the cross-citation records are grouped into two 5-year periods, namely, 2001–2005 and 2006–2010. Table 2 provides summary counts based on the proposed methodology. There are a total of 6,246 CPC subclass pairs that were first co-classified during the 2011–2013 period, and 92.99% of the fusions had cross-citations during the 2001–2010 period. The evidence supports the correlation between the stretching process and the final fusion event, which is in line with the findings of Caviggioli (2016), who used EPO data and the IPC system.

Figure 5 presents the distribution of the fusions across the levels of SD and CD. All the fusions having cross-citation(s) are assigned to four categories. Fusions in Category I experience cross-domain searches that involve low-level (not higher than 2) SD and CD. Fusions in Category II have high-level (higher than 2) SD and low-level CD cross-citations. Fusions in Category III have high-level SD and CD, and those in Category IV have low-level SD and high-level CD cross-citations. Table 3 shows the number of fusions in each category in the two time periods. The preliminary findings show that although the number of fusions with low-level SD and CD cross-citations is the highest among the four categories, nearly half of the fusions have high-level SD cross-citations. This phenomenon is more salient in the 2006–2010 period, during which significantly more fusions appear in Category III, i.e., with both high-level SD and CD, while fewer fusions appear in Category I. This finding implies that for a large part of technology fusions, the stretching between two technology domains is not limited to a certain narrow area of each domain.

To explore the dynamics of the cross-domain search process, fusions with cross-citations in both periods are extracted for in-depth analysis. Table 4 and Fig. 6 present how these fusions change their categories. Over 40% of the fusions in Category I in the 2001–2005 period changed to other categories in the 2006–2010 period, and over 25% of the fusions in Category II in the 2001–2005 period changed to Category III in the

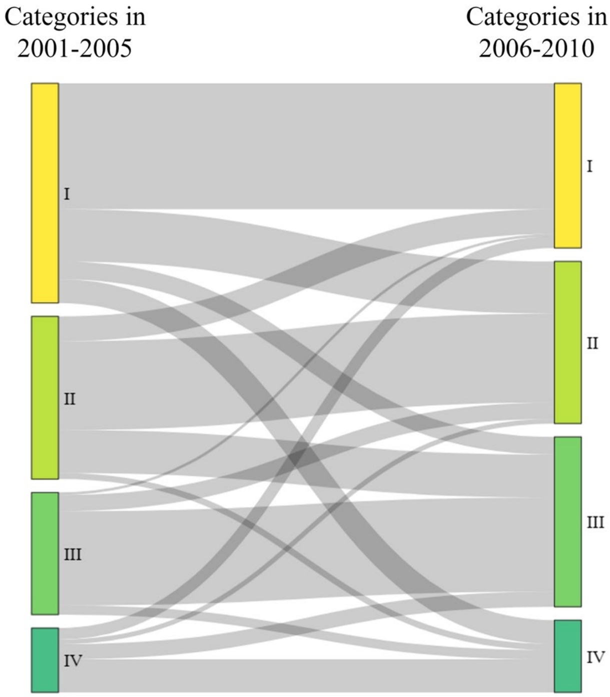


Fig. 6 The flow of category changes in fusions from the 2001–2005 period to the 2006–2010 period

2006–2010 period, which explains the decrease in fusions in Category I and the increase in fusions in Category III. Although some fusions have scope-narrowing cross-citations, for example, 221 fusions changed from Category II (2001–2005) to Category I (2006–2010), these net changes imply that the cross-domain search scope of technology fusions tends to expand.

The findings that address the extensive cross-domain search scope for technology fusion are in line with the assumptions of the proposed knowledge partition search model, which requires a broad search scope to not only identify attractive subspaces but also recognize and identify useful knowledge components. However, there are still numerous fusions with narrow stretching areas (e.g., Category I fusions) or even no cross-citations. In fact, in the search and recombination modeling literature, the initial knowledge of innovative agents is an important factor that shapes the search process (Liu & Ma, 2021; Puranam & Swamy, 2016). In the context of the proposed knowledge partition search model, if innovation agents are relatively knowledgeable about the technology domain, then the cross-domain search could be more targeted and thus may have a narrower scope.

To validate the above argument and justify the knowledge partition search model further, this paper considers the technological distance between merging domains. Theoretically, cross-domain searches would be more difficult if the merged technology domains were more distant because innovative agents could be more ignorant of distant technology domains. The results showed that 84.69% of the fused CPC subclasses were in different CPC sections (i.e., the first digits of the codes were different) and thus were generally in

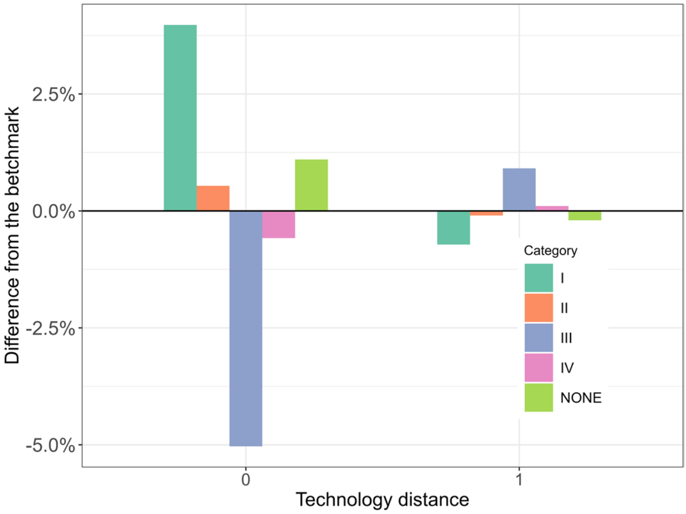


Fig. 7 The difference in category shares from the benchmark with different technology distances

distant areas. In this paper, a technology distance measure based on the CPC code is constructed as follows:

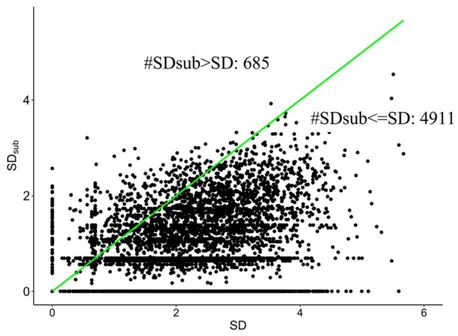
$$Tech\_dist = \begin{cases} 1 & differentCPCsections \\ 0 & sameCPCsection \end{cases} \tag{3}$$

where, if the merged technology domains are in the same CPC section, the distance between them is 0; if they are in different CPC sections, then the distance is 1.

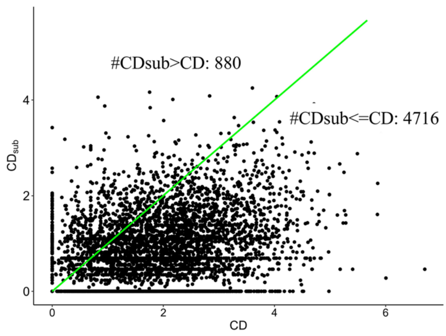
Focusing on the cross-citations in the 2006–2010 period and taking the share of each category in all the focal fusions as the benchmark, Fig. 7 presents the difference in the share of each category to the benchmark for fusions with different technology distances. The positive value of the difference in Fig. 7 indicates that the share is higher than the benchmark. The category "NONE" indicates that fusion had no cross-citation in the 2006–2010 period. Significant differences from the benchmark appear in fusions that merge technology domains in the same CPC section (distance is 0), which is associated with approximately 4% more fusions in Category I and 5% fewer fusions in Category III. That is, when merging technology domains in the same CPC section, agents may search across fewer subspaces, as well as fewer components, than when merging distant domains. This finding implies a more targeted cross-domain search and is in line with the argument about the influence of initial knowledge on the search process.

Forward-looking on technology fusions

In this subsection, the relationship between the characteristics of the stretching process and the fusions is explored further from the forward-looking perspective, i.e., considering the subsequent inventions of new hybrid technologies.



(a) Diversification at CPC main group level



(b) Diversification at CPC subgroup level

Fig. 8 Pre- and post-fusion SD and CD

Table 5 Descriptive statistics and correlations

Variables	Obs	Mean	Sd	Min	Max	1	2	3	4	5
1 Num_subinv	5596	3.30	5.39	1.00	136.00	1.00				
2 SD	5596	2.07	1.09	0.00	5.67	0.11	1.00			
3 CD	5596	1.74	1.09	0.00	6.70	0.18	0.39	1.00		
4 Num_cro	5596	2.14	1.17	0.00	5.69	0.20	0.65	0.71	1.00	
5 Tech_dis	5596	0.85	0.36	0.00	1.00	0.03	0.05	0.06	0.05	1.00

The diversification of subsequent inventions

Similar to the SD and CD measurements in Sect. "Methodology", the entropies at the CPC main group and subgroup levels of subsequent inventions in the 5-year window after the fusion event for each CPC pair could be calculated. For example, if one CPC pair, namely, X–Y was first co-classified in 2011, then the inventions that were applied during the 2011–2015 period in the X–Y area are collected. Following a method similar to that described in Sect. "Cross-domain search process measurement", the entropy-based diversification measures of the co-classified X–Y at the CPC main group (SD_{sub}) and subgroup level (CD_{sub}) in these subsequent inventions are calculated.

Considering only the fusions with cross-citation(s) in the 2006–2010 period, Fig. 8 compares the diversification of the search process and the diversification of subsequent inventions of each fusion at the CPC main group and subgroup levels. Interestingly, more than 80% of the fusions are distributed at the bottom right of each graph; i.e., the diversification of subsequent inventions tends to be lower than the diversification of the search process. This phenomenon is more salient in CPC main group level diversification (see Fig. 8a). This finding implies that actual fusion occurs among limited technology "points", despite extensive pre-fusion stretching attempts. Moreover, these findings are also in line with the basic idea of the proposed knowledge partition search model, which assumes that innovative agents search across the unknown landscape to identify (limited) valuable combinations.

Table 6 Results of the negative binomial and Poisson models

	Negative binominal				Poisson			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SD		− 0.006		0.002		− 0.002		0.004
		(0.014)		(0.015)		(0.009)		(0.009)
CD			0.066***	0.066***			0.064***	0.065***
			(0.016)	(0.016)			(0.009)	(0.009)
Num_cro	0.248***	0.251***	0.204***	0.203***	0.262***	0.264***	0.220***	0.217***
	(0.011)	(0.014)	(0.015)	(0.018)	(0.007)	(0.008)	(0.009)	(0.011)
Tech_dis	0.263***	0.263***	0.247***	0.247***	0.269***	0.269***	0.253***	0.252***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.026)	(0.026)	(0.026)	(0.026)
Year of fusion	Y	Y	Y	Y	Y	Y	Y	Y
CPC section	Y	Y	Y	Y	Y	Y	Y	Y
Constant	0.713***	0.718***	0.682***	0.680***	0.695***	0.697***	0.664***	0.660***
	(0.054)	(0.055)	(0.055)	(0.056)	(0.034)	(0.034)	(0.034)	(0.035)
Observations	5596	5596	5596	5596	5596	5596	5596	5596
Log Likelihood	− 12,386.090	− 12,385.990	− 12,377.540	− 12,377.530	− 16,156.160	− 16,156.120	− 16,132.110	− 16,131.990
Akaike Inf. Crit	24,796.170	24,797.980	24,781.080	24,783.060	32,336.310	32,338.250	32,290.220	32,291.980

Standard errors in parentheses. $*p<0.1$; $**p<0.05$; $***p<0.01$

The technological impact of technology fusions

A valuable new technology would attract significant attention from innovative agents. Understanding the relationship between the stretching process and the value of fusion would be of strategic importance because identifying or creating a new valuable technology earlier may confer considerable competitive advantage. How does the cross-domain search process influence the value of new hybrid technology?

To address this question, this study is interested in the relationship between the diversification of the search process (SD and CD) and the technological impact of fusions. Following Pezzoni et al. (2022), the technological impact of a fusion (Num_subinv) is measured as the number of subsequent inventions received in the 5-year window after the fusion event. Fusions that have cross-citation(s) in the 2006–2010 period are considered in the empirical analysis (including 5,596 samples). The number of cross-citations during the 2006–2010 period (Num_cro, in logarithm), the technological distance between the merged domains (Tech_dis), the year of the fusion event, and the CPC sections are controlled. Table 5 shows the descriptive statistics and correlations among the variables.

As the dependent variable consists of count data, both negative binomial and Poisson regressions are used for empirical analysis. Table 6 presents the regression results. The coefficient values of *CD* are significant and positive in Models 3, 4, 7, and 8, which implies that, compared to searching across different knowledge subspaces, the digging behavior for identifying refined knowledge components and combinations plays an important role in increasing the technological impact of fusions. Returning to the knowledge partition search model, the significant efficiency of digging behavior makes sense because efforts to recognize refined components for cross-domain combinations are necessary to identify valuable cross-domain combinations, i.e., new hybrid technologies with greater technological impact.

The coefficients of technological distance are significantly positive in all the models, which indicates that fusions of distant technology domains tend to have a greater technological impact. This finding is in line with the logic of recombination and search perspective, based on which some scholars measure the concept of “novelty” or “breakthrough” as the combination of distant knowledge components (Funk, 2014; Nemet & Johnson, 2012; Verhoeven et al., 2016).

Discussion

This study proposes a two-dimensional knowledge partition search model to understand the dynamics of stretching between distinct technology domains leading to technology fusions. Following the widely accepted methodology of technology fusion analysis, this paper is developed based on the availability of patent data and the hierarchical structure of CPC codes, which provides useful information for identifying technology domains, relationships, and trajectories. In particular, the hierarchical structure of the CPC matches well with the knowledge partition concept and thus facilitates revealing stretching dynamics by validating the telling knowledge partition search model.

In general, the stretching between distinct technology domains before an actual technology fusion event is not a “convergent” process. The empirical analysis shows that the search process of the majority of merged CPC pairs has a relatively high-level entropy at different CPC levels. A comparison of the cross-citation entropy between the 2001–2005 period and the 2006–2010 period reveals that although some samples experience a narrowing of the search scope, most of the changes are still at a similar level or even an increased level in terms of the diversity of the search scope. In other words, a large part of the search process is "divergent", in which innovative agents search across extending technology scopes to pursue the creation of technology fusion.

However, while the divergent search process is in line with the setting of the knowledge partition search model, it is also important to explain the difference in search processes among different technology fusions. Following the idea that initial knowledge held by innovative agents shapes the search process in the search and recombination literature, e.g., Puranam and Swamy (2016), this paper proposes that the pre-fusion search process of more knowledgeable innovative agents should be more targeted. The analysis that includes the technological distance between two merged technology domains supports this argument and further validates the proposed knowledge partition search model.

A broad cross-domain search scope does not naturally result in diversified fusion inventions. A comparison between pre-fusion and post-fusion entropy shows that for most of the fusions, the diversification of cross-citations is (much) greater than that of the subsequent inventions. In fact, this is also in line with the assumption of the knowledge partition search model, in which the useful combinations are distributed in limited positions.

In the same vein, extending the search scope cannot guarantee the technological impact of fusions. Compared to searching across different subspaces, digging is found to be positively associated with the technological impact of consequent technology fusions. This finding supports the knowledge partition search model setting, in which digging in subspaces is necessary for identifying useful knowledge components that may not have been recognized previously. This finding also helps to understand the source of technological breakthroughs or novelties, as some scholars emphasize long-distance recombination, while others believe that breakthroughs or novelties result from exploiting and refining current areas (Arts & Veugelers, 2015; Kaplan & Vakili, 2015; Nemet & Johnson, 2012). This paper suggests that long-distance recombination and knowledge refinement are not incompatible with each other. Rather, the knowledge refinement process facilitates actual long-distance recombination.

Conclusions

This paper reveals the general pattern of the cross-domain search process leading to technology fusion through empirically validating a knowledge partition search model. The findings show that, first, the cross-domain search process leading to technology fusion events is mainly "divergent"; that is, innovative agents gradually extend the search scope to pursue new hybrid technologies. Second, the general pattern of the cross-domain search process could be viewed as searching across a broad technology scope to identify limited valuable linking points within existing technology domains. However, it is notable that the specific search process for different technology fusions may also differ systematically. This paper suggests that cross-domain searches are more likely to be targeted when two technology domains are closer to each other, which is associated with the level of initial knowledge held by innovative agents.

In addition, compared to searching across a broader technology scope, digging in certain areas is more important for the generation of new high-impact hybrid technologies. This paper provides a new methodology for technology fusion analysis that has not been explored before; the methodology is closely associated with the debate between the "search deeply" and "search widely" concepts in the search and recombination literature. Although this paper does not focus on the firm-level problems, people could still borrow relevant ideas to explore the firm-level problems further. For example, future studies could empirically explore how the initial knowledge of firms influences their cross-domain search behavior.

This study has limitations that open opportunities for further research. First, the analysis in this paper is patent-based. Despite the benefit of the patent-based approach in technology fusion studies, patent data do not cover all the outcomes of innovative activity. Thus, this analysis may be useful only for sectors such as the pharmaceutical industry (Pezzoni et al., 2022), in which patents serve as proxies for inventions. Further research could use other datasets, for example, R&D investments or inventor collaboration, to conduct the analysis to reveal a more general pattern of the pre-fusion search process. In addition, further studies may also analyze the stretching process in or across specific domains for more purposeful policies. Second, this study measured the technological impact of technology fusion based on the count of reuse. There are also other measurements of technological impact, for example, the number of forward citations, which could be used in further studies. Moreover, further studies can also explore the linkage between the diffusion speed/distance of new hybrid technologies and the pre-fusion search process, which should be a fruitful direction for deepening the understanding of the spatiotemporal dynamics of technology diffusion.

Funding No funding was received for conducting this study.

Declarations

Conflict of interest The author has no conflicts of interest to disclose.

References

Ardito, L., Natalicchio, A., & Messeni Petruzzelli, A. (2023). Evidence on the determinants of the likelihood and speed of technological convergence: A knowledge search and recombination perspective in key enabling technologies. *IEEE Transactions on Engineering Management, 70*(4), 1442–1455.

Arthur, W. B. (2007). The structure of invention. *Research Policy, 36*(2), 274–287.

Arts, S., & Veugelers, R. (2015). Technology familiarity, recombinant novelty, and breakthrough invention. *Industrial and Corporate Change, 24*(6), 1215–1246.

Athreye, S., & Keeble, D. (2000). Technological convergence, globalisation and ownership in the UK computer industry. *Technovation, 20*(5), 227–245.

Bröring, S., & Leker, J. (2007). Industry convergence and its implications for the front end of innovation: A problem of absorptive capacity. *Creativity and Innovation Management, 16*(2), 165–175.

Castaldi, C., Frenken, K., & Los, B. (2014). Related variety, unrelated variety and technological breakthroughs: An analysis of us state-level patenting. *Regional Studies, 49*(5), 767–781.

Cavalheiro, G. M. D. C., Joia, L. A., & Van Veenstra, A. F. (2016). Examining the trajectory of a standard for patent classification: An institutional account of a technical cooperation between EPO and USPTO. *Technology in Society, 46*, 10–17.

Caviggioli, F. (2016). Technology fusion: Identification and analysis of the drivers of technology convergence using patent data. *Technovation, 55–56*, 22–32.

Curran, C.-S., & Leker, J. (2011). Patent indicators for monitoring convergence – examples from NFF and ICT. *Technological Forecasting and Social Change, 78*(2), 256–273.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science, 47*(1), 117–132.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy, 30*(7), 1019–1039.

Funk, R. J. (2014). Making the most of where you are: Geography, networks, and innovation in organizations. *Academy of Management Journal, 57*(1), 193–222.
Galunic, D. C., & Rodan, S. (1998). Resource recombinations in the firm: Knowledge structures and the potential for schumpeterian innovation. *Strategic Management Journal, 19*(12), 1193–1201.
Hötte, K., Pichler, A., & Lafond, F. (2021). The rise of science in low-carbon energy technologies. *Renewable and Sustainable Energy Reviews, 139*, 110654.
Jaffe, A. B., & Trajtenberg, M. (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences of the United States of America, 93*(23), 12671–12677.
Jeong, S., & Lee, S. (2015). What drives technology convergence? Exploring the influence of technological and resource allocation contexts. *Journal of Engineering and Technology Management, 36*, 78–96.
Kangas, H. L., Ollikka, K., Ahola, J., & Kim, Y. (2021). Digitalisation in wind and solar power technologies. *Renewable and Sustainable Energy Reviews, 150*, 111356.
Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal, 36*(10), 1435–1457.
Karvonen, M., & Kässi, T. (2013). Patent citations as a tool for analysing the early stages of convergence. *Technological Forecasting and Social Change, 80*(6), 1094–1107.
Keijl, S., Gilsing, V. A., Knoben, J., & Duysters, G. (2016). The two faces of inventions: The relationship between recombination and impact in pharmaceutical biotechnology. *Research Policy, 45*(5), 1061–1074.
Kim, E., Cho, Y., & Kim, W. (2014). Dynamic patterns of technological convergence in printed electronics technologies: Patent citation network. *Scientometrics, 98*(2), 975–998.
Kim, M.-S., & Kim, C. (2012). On a patent analysis method for technological convergence. *Procedia - Social and Behavioral Sciences, 40*, 657–663.
Klarin, A., Suseno, Y., & Lajom, J. A. L. (2023). Systematic literature review of convergence: A systems perspective and re-evaluation of the convergence process. *IEEE Transactions on Engineering Management, 70*(4), 1531–1543.
Knudsen, T., & Srikanth, K. (2014). Coordinated exploration: Organizing joint search by multiple specialists to overcome mutual confusion and joint myopia. *Administrative Science Quarterly, 59*(3), 409–441.
Kwon, O., An, Y., Kim, M., & Lee, C. (2020). Anticipating technology-driven industry convergence: Evidence from large-scale patent analysis. *Technology Analysis & Strategic Management, 32*(4), 363–378.
Lee, C., Hong, S., & Kim, J. (2021). Anticipating multi-technology convergence: A machine learning approach using patent information. *Scientometrics, 126*(3), 1867–1896.
Liu, J., & Ma, T. (2021). How the timing of cooperation affects innovation outcomes: An agent-based model of postponing the coupled search process. *SIMULATION, 97*(11), 739–759.
Nakamura, H., Suzuki, S., Sakata, I., & Kajikawa, Y. (2015). Knowledge combination modeling: The measurement of knowledge similarity between different technological domains. *Technological Forecasting and Social Change, 94*, 187–201.
Nemet, G. F., & Johnson, E. (2012). Do important inventions benefit from knowledge originating in other technological domains? *Research Policy, 41*(1), 190–200.
No, H. J., & Park, Y. (2010). Trajectory patterns of technology fusion: Trend analysis and taxonomical grouping in nanobiotechnology. *Technological Forecasting and Social Change, 77*(1), 63–75.
Oh, S., Choi, J., Ko, N., & Yoon, J. (2020). Predicting product development directions for new product planning using patent classification-based link prediction. *Scientometrics, 125*(3), 1833–1876.
Park, H., & Yoon, J. (2014). Assessing coreness and intermediarity of technology sectors using patent co-classification analysis: The case of Korean national R&D. *Scientometrics, 98*(2), 853–890.
Pezzoni, M., Veugelers, R., & Visentin, F. (2022). How fast is this novel technology going to be a hit? Antecedents predicting follow-on inventions. *Research Policy, 51*(3), 104454.
Preschitschek, N., Niemann, H., Leker, J., & G Moehrle, M. (2013). Anticipating industry convergence: Semantic analyses vs IPC co-classification analyses of patents. *Foresight, 15*(6), 446–464.
Puranam, P., & Swamy, M. (2016). How initial representations shape coupled learning processes. *Organization Science, 27*(2), 323–335.
Rosenberg, N. (1963). Technological change in the machine tool industry, 1840–1910. *The Journal of Economic History, 23*(4), 414–443.
Samuelson, L. (2004). Modeling knowledge in economic analysis. *Journal of Economic Literature, 42*(2), 367–403.
Sick, N., & Bröring, S. (2022). Exploring the research landscape of convergence from a TIM perspective: A review and research agenda. *Technological Forecasting and Social Change, 175*, 121321.
Singh, A., Triulzi, G., & Magee, C. L. (2021). Technological improvement rate predictions for all technologies: Use of patent data and an extended domain description. *Research Policy, 50*(9), 104294.

Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy, 45*(3), 707–723.

Xiao, T., Makhija, M., & Karim, S. (2022). A knowledge recombination perspective of innovation: Review and new research directions. *Journal of Management, 48*(6), 1724–1777.