ORGANIZING FOR INNOVATION: HOW TEAM CONFIGURATIONS VARY WITH MODULARITY AND BREADTH OF APPLICATION

Keyvan Vakili

London Business School, Strategy and Entrepreneurship Department

Sussex Place, London, United Kingdom, NW1 4SA

kvakili@london.edu

Sarah Kaplan

Rotman School of Management, University of Toronto

105 St. George St., Toronto, ON, Canada, M5S 3E6

sarah.kaplan@rotman.utoronto.ca

Keywords: breakthrough innovation, team configuration, recombination, modularity, general purpose technologies, knowledge, patents

Acknowledgement: This work is partially supported by the Canadian Social Sciences and Humanities Research Council under grant #410-2010-0219. All errors and omissions remain our own.

ORGANIZING FOR INNOVATION: HOW TEAM CONFIGURATIONS VARY WITH MODULARITY AND BREADTH OF APPLICATION

ABSTRACT

While innovation has increasingly become a collaborative effort, there is little consensus in research about what types of team configurations might be the most useful for creating breakthrough innovations. Do teams need to include inventors with knowledge breadth for recombination or do they need inventors with knowledge depth for identifying anomalies? Do teams need overlapping knowledge to integrate insights from diverse areas or does this redundancy hamper innovation by creating inefficiencies? In this paper, we suggest that the answers to these questions depend on the characteristics of the technologies, which explains why prior evidence based on single domains or that aggregates all technologies have yielded inconclusive findings. Focusing on the degree of modularity and the breadth of application in patent data, we find that differing team configurations are associated with different technological domains.

INTRODUCTION

Innovation has increasingly become a collaborative effort (Wuchty et al, 2007). As knowledge domains advance and individuals become more specialized, inventors need to engage in more collaborative efforts to advance the knowledge frontier (Jones, 2009; Agrawal, Goldfarb, and Teodoridis, 2016). Data from United States Patent and Trademark Office (USPTO) patents shows that the average team size of innovative teams grew from 1.7 in 1975 to more than 2.5 in 2010. Other research suggests that teams are better at filtering out low-quality ideas and thus are likely to produce more impactful innovations than lone inventors (Singh and Fleming, 2010). Accordingly, a long stream of research has argued that there is an important relationship between the configuration of innovative teams and their innovation outcomes (Perry-Smith and Shalley, 2014; Bercovitz and Feldman, 2011; Dahlin, Weingart, and Hinds, 2005; Kogut and Zander, 1992; Taylor and Greve, 2006).

A challenge in this literature is that scholars have not addressed differences that may arise across different technological domains and this may be a source for the contradictory findings about the nature of this relationship. For example, while a large body of research claims that recombinations of distant and diverse knowledge areas are associated with innovation novelty (Fleming, 2001; Gittelman and Kogut, 2003), others have reported that depth and specialization is more important (Kaplan and Vakili, 2015, Teodoridis, Bikard, and Vakili, 2019). Similarly, whereas some have shown a positive relationship between knowledge breadth and creative output (Hargadon and Sutton, 1997; Burt, 2004; Uzzi and Spiro, 2005), others have found knowledge breadth to stymie such creativity (Leahey 2007; Leahey, Beckman, and Stanko 2017).

One possible explanation of these contradictory findings is that they are based on empirical evidence from different technological domains. From the days of Joan Woodward (1958, 1965), we have known that organizational form should depend on the types of technologies the organization pursues. Yet, prior research on team configurations for innovation has not taken up this insight in their studies. Instead, scholars either study innovations drawn from a wide set of technological domains, often the whole population of patents (e.g., Singh and Fleming, 2010; Arts and Veugelers, 2014; Fleming et al 2017) and simply control for technological differences, or they focus on a single technological domain such as biotechnology

(Phene et al, 2006), the chemical industry (Rosenkopf and Nerkar, 2001) or carbon nanotubes (Kaplan and Vakili 2015). In both cases, the claims about organizational form are generally agnostic to the technological domain in which the innovation takes place. That is, average effects in a broad range of technological domains might disguise opposing effects within specific domains. Alternatively, specific effects identified in one technological domain might not be applicable elsewhere. In most of these studies, any acknowledgement of the boundary conditions that technological domains might impose on the findings usually appears in a discussion of limitations but is not central to the theorizing.

We argue that abstracting away from the role of technological domain can lead to theoretical and empirical confusion as well as misleading recommendations for practice, and may be part of the explanation for the conflicting findings in the field to date. The insights and practical implications drawn from one domain could potentially lead to undesirable outcomes if applied in another domain with different underlying characteristics. In this paper, we take a step towards addressing these issues by comparing the relationship between team configuration and innovation outcomes in four different technological domains. In this cross-domain analysis, we focus specifically on differences in the degree of modularity and the breadth of application, each of which relate to important streams of research on technological characteristics.

Modularity is defined as the extent to which a technology can be decomposed to a set of components with standardized interfaces between them (Langlois, 2002; Brusoni and Prencipe, 2001). The fields of electronics and computer programming are among those highly associated with modular technological design. In contrast, non-modular technologies such as chemistry or biology usually cannot be decomposed into their sub-components and involve complex mapping between their physical and functional elements (Ulrich, 1995). The literature on modularity suggests that modularity can support a more efficient division of labor among team members by removing the need for knowledge redundancies. It facilitates integration through standardization. Hence, for modular technologies, teams with lower levels of overlap can achieve higher levels of knowledge diversity while avoiding the inefficiency of redundancy, hence producing more novel innovations. By corollary, for non-modular technologies, knowledge overlap will be needed to integrate insights.

The breadth of application of a technology is a second characteristic that can moderate the performance of innovative teams with varying knowledge compositions. Technologies vary substantially in their range of applications. At one extreme, technologies such as electricity and computer chips can find applications across a very large range of domains (Helpman and Trajtenberg, 1994). At the other extreme, some technologies such as coronary stents and electrocardiography are developed for narrow sets of applications. We argue that where there is a greater breadth of application of a technology, novel innovative outcomes will be more associated with innovative teams that have a wider span of knowledge and less knowledge redundancy. In contrast, depth of knowledge may be associated with the development of novel innovations based on narrow-purpose technologies.

Importantly, technologies can vary on both dimensions. We exploit this variation in a cross-domain analysis of patents in four different technological domains—magnetic resonance imaging (MRI), radio frequency identification (RFID), stem cells, and nanotubes, each of which represent one box in a two-by-two of modularity and breadth of application (Table 1). Whereas the first two domains are electronics-based and highly modular, the latter two are chemistry-based and less modular. Also, while RFID and nanotubes have found applications across a wide and diverse range of areas, MRI and stem cells have relatively narrower set of applications. By examining the effect of team configuration on innovative outcomes in these four technological domains, we can shed light on the contingent organization designs necessary for innovation.

-- Insert Table 1 about here --

Using two different measures of innovative outcomes—patents that represent knowledge breakthroughs and those that represent economic breakthroughs—we look at how different team configurations might be associated with each within different technological domains. To characterize the configuration of innovative teams, we examine the knowledge breadth and depth of the main inventors of a team as well as the overlap in their knowledge. Past research has highlighted the tradeoffs associated with knowledge depth versus breadth at individual (Leahey, Beckman, and Stanko, 2016), technological (Kaplan and Vakili, 2015), team (Bercovitz and Feldman, 2011), and organizational (Ahuja and Lampert, 2001) levels. Whereas a breadth of knowledge provides access to more diverse and distant knowledge bases to be used in knowledge recombination (Hargadon & Sutton 1997; Audia & Goncalo 2007), the depth of knowledge can help with understanding the foundations and major gaps in a domain (Kuhn 1962; Taylor & Greve 2006, Weisberg, 1999). Moreover, we focus on the role of overlap in the knowledge of team members as a bridge that facilitates the integration of their individual knowledge stocks. Past research suggests that knowledge overlap reduces communication costs between team members and can lead to more effective teamwork (Dahlin et al, 2005; Dougherty, 1992). On the other hand, overlap may also lead to knowledge redundancies which, in turn, can lower the level of recombinant opportunity in the team or create inefficiencies (Burt, 2004).

Based on our analyses, we find that different team configurations are associated with innovative outcomes in different technological domains. While we cannot make causal claims, we find that in highly modular technological domains, the production of novel breakthroughs is associated with teams with less overlap between their inventors' knowledge scope, while in non-modular domains, it is associated with teams in which the most experienced inventor can act as the knowledge integrator. Further, novel breakthroughs in technologies with broad applications are associated with teams in which the second most experienced inventor has wide knowledge breadth, presumably to seek out a wide range of applications. In contrast, novel breakthroughs in narrow technologies are associated with teams in which second inventor has more knowledge depth, presumably to seek out anomalies. Using our cross-domain research design, we then empirically show and discuss how the interaction of modularity and application breadth moderates the associations of different team configurations with the chance of producing novel breakthroughs.

We also find that the team configurations that are associated with the production of novel breakthroughs are not always aligned with those that are associated with the production of economic breakthroughs. Our estimations suggest that the knowledge breadth of the first inventor is one factor that is positively, and consistently across all four domains, associated with the likelihood of producing economic breakthroughs. This may suggest that knowledge breadth is better understood as a proxy for the broader networks and social connections that inventor might possess that would lead to greater diffusion and impact.

Our analysis gives support to the intuition that team design might be contingent on the technological domain. We provide theoretical arguments that explain the observed associations and also shed light on some of the empirical and theoretical inconsistencies in prior research on organizational design for innovation.

THEORETICAL BACKGROUND

Innovative Team Configuration

Past research on organization design in innovation points to two distinct factors that are associated with innovation outcomes: knowledge breadth versus depth and the level of overlap in knowledge stock. The former relates to the knowledge that enters into the innovation process, whereas the latter largely relates to the process of knowledge integration.

There is substantial research on the degree to which knowledge specialization or knowledge breadth is required for innovation, but the arguments are divided. On the one hand, scholars highlight the importance of knowledge diversity for developing novel and impactful innovations (Powell, Koput, and Smith-Doerr, 1996; Jeppesen and Lakhani, 2010; Leiponen and Helfat, 2011; March, 1991). The argument relies on the idea that innovation is a recombination process. New ideas are essentially combinations of previously disconnected ideas. Hence, more diversity in knowledge input can lead to more novel knowledge recombinations and consequently more impactful innovations (Hargadon and Sutton, 1997; Burt, 2004; Fleming, 2001; Schilling and Green, 2011). To maximize their impact, teams and organizations should increase the scope of their knowledge by investing in interdisciplinary projects, hiring people from diverse knowledge backgrounds, or forming alliances with more diverse partners (Sampson, 2007; Schilling and Green, 2011; Østergaard, Timmermans, and Kristinsson, 2011; Smith, Collins, and Clark, 2005).

On the other hand, research highlights the importance of knowledge depth. This research argues that specialists have a better understanding of the fundamental gaps in their domain of specialty (Weisberg, 1999), can absorb and use the knowledge at the frontier more effectively (Jones, 2009; Teodoridis, Bikard and Vakili, 2019), and have superior domain-specific memory and problem-solving skills (Larkin et al., 1980; Sweller, Mawer, and Ward, 1983). From this standpoint, to teams and organizations should invest in domain-specific expertise to maximize their impact.

The effect of knowledge overlap for innovation output is also similarly debated in the literature. On the one hand, knowledge overlap helps with facilitating knowledge integration. A lack of common language can increase the communication costs between team members and lead to undesirable frictions through the process (Krauss and Fussell, 1990; Cramton, 2001). This is particularly the case when there is a need for cross-boundary knowledge integration. The coordination costs are higher when individuals have different knowledge backgrounds and when links between knowledge domains are less established. The existence of overlapping knowledge bases can thus facilitate the knowledge integration process at the team level. The shared knowledge base helps individuals find a common language and find the linkages between the disconnected parts of their knowledge.

On the other hand, past research also points out the undesirable redundancies in knowledge base for innovation (Burt, 2004). All things being equal, an increase in overlap in individuals' knowledge bases comes at the expense of a decrease in the total diversity of knowledge at the team level. Hence, while knowledge overlap can facilitate the knowledge integration process, it can nevertheless lead to lower diversity in knowledge input.

A few scholars have pointed to contingencies in these relationships. Leahey et al. (2017) show that scientists who engage in more interdisciplinary projects (i.e., rely on wider knowledge breadth) are more likely to produce very impactful innovations, but they experience a decline in their productivity. In contrast, they show that specialization is associated with higher innovation rate, but lower impact. Others have explored the non-linear effects of knowledge breadth and overlap, suggesting that while they are both necessary for the production of novel and impactful innovations, too much can eventually lead to negative effects (Katz, 1982; Berman, Down, and

Hill, 2002).

Inspired by these insights, in this paper, we explore one direction for addressing the contrasting findings that predominate in the literature by using a cross-domain analysis to explore the technological contingencies that can influence the benefits and costs associated with these aspects of team configuration. We specifically discuss how differences in the characteristics of a technological domain would be associated with different team configurations for innovation.

Technological Domain

Recent research on organization and team design for innovation has largely treated technology characteristics as a boundary condition rather than a moderating factor deserving its own theoretical and empirical investigation. The idea that technology characteristics might be associated with different team configurations has long roots in innovation research. Going back as early as Woodward (1958), scholars have highlighted the importance of the match between technology and organization design. Woodward (1958) specifically distinguishes between three types of technologies—large batch and mass production, unit and batch production, and continuous processing—and argues that while centralized bureaucratic decision-making structures are more effective for the former technology type, organizations with decentralized decision-making structure are better fit to the latter two types of technology. These insights were foundational to the formation of contingency theory, which argued for a match between technology type and the level of task variability as well (Thompson, 1967; Perrow, 1967).

However, when it comes to work on team configurations for innovation, the specific recommendations of contingency theory have not been taken up. Scholars either study innovations drawn from a wide set of technological domains and make claims about average effects while only for technological differences (e.g., Singh and Fleming, 2010; Arts and Veugelers, 2014; Fleming et al 2017), or they focus on a single technological domain in the hopes that the insights might be applicable elsewhere (Phene et al, 2006; Rosenkopf and Nerkar, 2001; Kaplan and Vakili 2015; Bikard, Vakili and Teodoridis, 2018). In this paper, we build on this legacy of insights from contingency theory and demonstrate that some team configurations are more likely to be associated with producing novel and impactful innovations depending on different technological characteristics. We specially focus on two characteristics of a technological domain: application breadth and modularity.

Application breadth refers to the domains in which a technology can be used. At the extreme, some technologies may be applied across many diverse and distant areas. Electricity, wheels, plastics, internet, artificial intelligence, and nanotechnology are some examples of technologies with broad applications. There is relatively less research on the relationship between the application breadth of a technology and organizational design elements. What type of team configuration is more likely to produce cognitively novel and economically impactful innovations within the context of broadly-applied technologies? Most broadly-applied technologies have their roots in a specific technological domain, but innovations based on them usually bridge the technology to new application domains. For example, nanotubes were first invented by a physicist, Sumio Iijima, at the NEC corporation. However, most novel innovations based on nanotubes involve exploiting a specific mechanical, chemical, or electrical property of nanotubes in a new application domain. Hence, when the technology is broadly-applied, we expect inventors' knowledge breadth to be essential for discovering new application domains for the technology and producing cognitively novel innovations. Moreover, too much overlap between the top inventors in such domains can limit the team's ability to explore new application domains. In contrast, when technologies are only narrowly-applied, we expect inventors' knowledge breadth to be less helpful in developing novel and impactful innovations.

Technological modularity generally refers to the extent to which a technology can be decomposed into distinct components each responsible for a specific functionality. In modular technologies, there is usually a one-to-one mapping between physical components and functional elements (Ulrich, 1995). This is highly related to the original distinction that Woodward (1958) drew between batch and process manufacturing. There has been a growing interest in the implications of modularity for organization design. The research, however, has predominantly focused on product-level modularity and its impact on organizational design choices such as vertical integration (Brusoni and Prencipe, 2001; Langlois, 2002) and hierarchical coordination (Sanchez and Mahoney, 1996). At the team level, we expect technological modularity to substitute for knowledge overlap in the integration process. A one-to-one mapping between components and functionalities allows individual inventors to focus on components, relying on the standardized interfaces between components for team-level integration. While there is usually a need for a lead inventor to design the system at the top level, the lead does not need to have extensive knowledge about each component and its intricacies. Thus, in modular technologies, we expect novel and impactful innovations to be produced by teams with lower knowledge overlap between the lead inventor and others. What is less known is the degree to which this intuition holds when the technology can be applied broadly or only narrowly.

In contrast, for non-modular technologies, we expect the lead inventor to play a more significant role in team-level knowledge integration. While most research on organizing for modularity is at the organizational level, not the team level, that research implies that in non-modular domains, knowledge overlap plays a crucial role in knowledge integration (Grant, 1999). It is hence important for the lead inventor to have a good understanding of other team members' knowledge to be able to communicate effectively with them and to facilitate integration between them. In short, in non-modular technologies, we expect teams in which the lead inventor has more knowledge breadth and more overlap with other inventors to be more likely to produce novel and impactful innovation.

Our goal in this paper is to explore how the team configurations that produce breakthrough innovation might differ based on the nature of the technological domain in which they operate—modular vs. non-modular and broad vs. narrow application. We use this cross-domain analysis to shed light on the contrasting findings in extant research about whether breadth or depth of knowledge matters and how much knowledge overlap would be required for an innovation team. In our analysis, we observe different team configurations are associated with innovative outcomes depending on the technological domain. Indeed, we show that one cannot draw a conclusion about team configurations for modular or non-modular technologies without considering also the breadth of application and vice versa.

METHODOLOGY

Data Sample

We use evidence from teams listed in patents from the US Patent and Trademark Office for this analysis. Patents, despite their shortcomings as a measure of innovation (Mansfield, 1986), provide evidence of successful innovations and document clearly the team members who participated in the innovation process. Patent data also allow us to measure the depth and breadth of experience of each listed inventor by tracing their other granted patents. To explore the role of modularity and technological breadth in moderating the association of team configuration on innovation output, we need patents from a theoretical sample of four technological domains that represent each unique combination of modularity and breadth of application. In other words, the analysis requires at least one technological domain that is modular and applied broadly, one that is non-modular and applied broadly, one that is modular and applied narrowly, and one that is both non-modular and applied narrowly. Moreover, we should be able to trace the history of each technological domain for a long enough time using patent data: technologies should not be so recent as to have few patents and not so old as to have a bulk of patents prior to 1976 which is the cutoff for the digitization of patent data. Finally, we need to focus on technological domains that have a roughly identifiable boundary to be able to identify the set of innovations that belong to each domain. In practice, we need to be able to specify the patented innovation in each domain using a set of keywords without having too many false positives or false negatives.

Given these requirements, we identified a comprehensive sample of patents in four technological domains: Magnetic Resonance Imaging (MRI), Radio Frequency Identification (RFID), stem cell, and nanotubes. MRI is a (modular and narrowly-applied) medical imaging technology used to scan internal organs and physiological processes of humans and animals. RFID is a (modular and broadly-applied) technology based on electromagnetic fields used to identify and track pre-programmed tags in a certain physical range. Stem cells are a (non-modular and narrowly-applied) technology based in undifferentiated cells that have the capability to develop into specialized cell types. Nanotubes are a (non-modular and broadly-applied) technology made of cylindrical materials made of carbon molecules with nanometer scale diameters and special properties such as superconductivity and high levels of elasticity. These four technologies vary substantially across the two technological dimensions of interest.

Our sample consists of all patents granted by the United States Patent Office (USPTO) in each domain between 1970 and 2010. For patents prior to 1976, we hand searched the patent database because fully-digitized records were not available. The year 2010 was used because of our reliance on the NBER Patent Dataset and the Harvard Patent Dataverse (Lai et al., 2015) for complementary data, both of which only cover patents granted until 2010. Moreover, ending our sample in 2010 allows us to trace the impact of the patents on follow-on innovations (forward citations) over subsequent years.

We used several complementary methods to identify patents in each technological domain. We first searched for all patents that mention any of these technologies or their variants in their title, abstract, or claims. In the case of RFID, we further included all the patents in the technological class 340/13.26 designated by the USPTO for RFID patents. In the case of nanotubes, we complemented our search by selecting all patents in the cross-reference subclasses 977/735-752 assigned by the USPTO retroactively to nanotube patents. We also used Derwent's technological classification to select all patents in classes B05-U, C05- U, E05-U, E31-U02, L02-H04B, U21-C01T, X12-D02C2D, X12-D07E2A, X12-E03D, X16-E06A1A, all of which are related to nanotubes. In the case of stem cells, we also added all the patents that were listed on the StemCellPatents.com website and granted before 2010. MRI does not have its own patent classification code. Occasionally multiple patents may be granted to protect the same invention. These patents are usually recognized as a single patent family and have very similar abstracts. For each patent family, we only use the patent with the earliest application year in our sample.

Table 2 lists the set of search terms used for each technology, the number of patents retrieved from each search, and the total number of patents resulting from all searches for each technology after collapsing all patents belonging to the same family into one. Our final sample includes 9,230 MRI patents, 3,521 RFID patents, 1,775 stem cell patents, and 2,384 nanotube patents. There are small overlaps between the four groups of patents: 4 patents belong to both

RFID and nanotubes categories; 6 patents belong to both MRI and RFID; 14 patents are classified in both MRI and nanotubes; 7 patents are classified in both MRI and stem cells; and, 3 patents belong to both stem cells and nanotubes. The overlaps are small, and all results are robust to excluding the patents belonging to more than one category.

-- Table 2 about here --

Technological Modularity and Application Breadth

Importantly for the purposes of this analysis, each of these four technological domains are positioned differently with respect to the two technological characteristics of interest: modularity and application breadth (Table 1 above).

Breadth of application. Whereas innovations in RFID and nanotubes domains have a wide range of applications, MRI and stem cells innovations are much narrower in their application. Variants of RFID technology have found applications in numerous domains such as physical tracking of material through the supply chain, item-level monitoring in the manufacturing process, queue optimization in hospitals and amusement parks, race timing, production of robbery-proof chips for casinos, library management systems, interactive marketing, and attendee tracking in large conferences. Similarly, nanotubes have found applications from producing new semiconductor materials to developing reinforced composite used in golf balls to improving the performance of fuels. MRI and stem cells have relatively more limited application areas. To date, stem cells have been only used for specific therapeutic applications in animals and humans. Until the 1970s, MRI technology was used mostly for chemical and physical analysis. However, the advent of superconductors capable of producing strong magnetic fields made it possible to use MRI technology to scan human body parts for the first time in 1977. Since then, the technology has predominantly been used for non-invasive medical imaging purposes.

We can infer the application breadth of each technological domain based on the characteristics of their patents. Broadly-applied technologies are pervasive and are subject to continuous growth in their domains of applications (Hall and Trajtenberg, 2004). Observing the subsequent patents that cite a focal patent, we can infer that technologies with broader applications would be cited by patents in a wider range of technological classes (Shea et al., 2011), where technological classes are assigned by the patent office to demarcate technological domains. We can also measure the growth in technological classes of citing patents as a proxy for the growth in application domains of each technology. Table 3 shows these two measures for patents of each selected technology. RFID and nanotube patents (broadly-applied technologies) are on average cited by subsequent patents in 6.2 and 6.8 unique 3-digit technological classes. In contrast, MRI and stem cell patents (narrow purpose technologies) are cited by patents in 3.4 and 3.9 unique classes, respectively. The t-test analysis shows that the differences between the average number of citing classes of RFID and nanotube patents and that of MRI and stem cell patents are significant at the 99% level. Moreover, whereas the number of unique technological classes of patents citing RFID and nanotubes patents grew during the sample period by 24% and 18%, respectively, it is 5% and 0% for MRI and stem cell patents respectively.

-- Table 3 about here --

While, in line with our theoretical conceptualization, the majority of patents in the nanotubes and RFID domains involve applying these technologies to various applications, a small number of patents in each domain relate to inventions focused on developing the technologies themselves. Specifically, 128 patents in the RFID domain and 115 patents in the nanotubes domain are focused on developing or advancing RFID technology and nanotubes respectively. To have better alignment between our theoretical conceptualization and empirical analysis, we exclude these patents in our main results. Nevertheless, in the appendix we provide evidence showing that our results are robust to including the RFID and nanotube patents that do not share the breadth of application of their domain (Table A9 in the appendix).

In our empirical analysis, we use two independent variables associated with these two characteristics of technologies. Since our goal is to highlight the moderating role of technological domain, we use the variances at the technological domain level (instead of patent level) in our estimations. The $Modular$ variable specifies whether a technology is modular or not and is equal to 1 for MRI and RFID patents, and 0 for stem cell and nanotube patents. $Broadly\_applied$ variable specifies whether a technology is broadly-applied or not and is equal to 1 for RFID and nanotube patents and is 0 otherwise. We acknowledge that not all patents in a technological domain share the typical characteristic of the domain. However, our goal is to show how team configurations behind breakthrough innovations vary at the technological domain. Therefore, our simplifying assumption is aligned with the purpose of our analysis.

We also perform our estimations at the technological domain level to estimate how the relationship between the knowledge composition of innovative teams and innovation outcomes varies across each possible combination of these two characteristics of the technology: modular and broad application (RFID), modular and narrow application (MRI), non-modular and broad application (nanotubes), and non-modular and narrow application (stem cell).

Modularity. The technologies can also be categorized based on their level of modularity. MRI and RFID are relatively more modular than stem cells and nanotubes. MRI and RFID both have their roots in electronics. A standard MRI device is composed of an external magnetic field, a set of gradient coils, RF equipment, power supply, display unit, a computing unit and a set of computer programs to analyze and display the data collected from the imaging process. Each of these components may be divided into various sub-components. For example, the computing unit itself is built of a processing unit, memory chips, graphic cards, and input devices. The interactions between these components are highly standardized. An improvement in the code that analyzes the input from the imaging unit does not require a change in the whole system. The individuals who write the code to analyze the output of the imaging unit do not need to know how the coils move. Similarly, those who write the code for the display unit do not need to understand the structure of the output data from the imaging unit. It is relatively straightforward to map the functional elements (imaging, data processing, output display) to physical elements. The interface between the units follow established protocols of electronic data communication.

RFID technology is similarly composed of a set of standard components. The system includes tags, a reader, antenna, and a computing unit with application software. The tags listen for signals sent by a reader. When they receive a query signal, they respond by sending their unique id back to the reader. Both parts rely on communication chips and antenna systems and are composed of more sub-components. Again, the interaction between these components are highly standardized. In fact, there are two large standards bodies, ISO RFID and EPCglobal, that specify and supervise the standardization of RFID systems and elements. The standards range from coding to tag data and tag-reader linkages. Similar to MRI technology, it is straightforward to map functional elements of the technology to its physical elements.

In contrast, stem cell and nanotubes have their roots in biology and chemistry, respectively, and have much less modular designs. They are not produced by putting a set of sub-components together and it is difficult to map their functions to specific physical elements. Stem cells are generally extracted from animal or human tissues and then grown in laboratory, a process known as cell culture. None of the constituent components of stem cells such as protein, DNA, and RNA can function in the absence of other components. Relatedly, scientists have not yet managed to extract and work on these components separately and then put them together to manufacture a stem cell with specific desirable properties. The same is largely true for nanotubes. They are produced through chemical or electrochemical processes and, depending on the production process, they may end up with different properties such as superconductivity or physical resistance. While it is possible to produce nanotubes separately and then combine them with other materials, nanotubes themselves cannot be manufactured by putting their constituent elements (i.e., carbon molecules) together; at least not yet. The process of producing both nanotubes and stem cells are highly complex and prone to inaccuracies. For these reasons, it is difficult to mass produce them and their production is generally limited to certain laboratories.

Innovative Team's Knowledge Composition

Our analysis focuses on how team configurations might differ for each of these different technological domains. A fundamental question in organization design is the deployment of specialists and generalists and the degree of coordination required between people (Stan and Puranam, 2016; Brusoni, Prencipe, and Pavitt, 2001). Specialists have deep knowledge in one domain: they are needed when there are economies of scale for knowledge and integration across multiple domains is less essential. Generalists have broad knowledge across multiple domains but may not be as deep. They are needed when cross-boundary knowledge integration is essential for creating and capturing value.

In the case of inventive teams, we are concerned about the knowledge depth and breadth of the main inventors and the overlap in their knowledge. For our analyses, we focus on the two inventors of each team with the greatest amount of experience during the 5 years prior to their focal innovation (hereafter 5-year patenting experience) (not the first two listed in the document). The median innovative team in our sample has only two inventors. The average team size is 2.6. We gain very little information by expanding beyond the first two team members. For teams that have more than 2 inventors, the median five-year patenting experience of the third, fourth, and fifth inventor (ranked by their 5-year experience) is zero. Moreover, their average five-year patenting experience is below two patents. Finally, their knowledge scope has more than 80% overlap with the lead inventor. Hence, the majority of the third, fourth and fifth inventors in our sample have little inventive experience to contribute to the team. Nevertheless, we control for their experience and the total number inventors in a team in all specifications. While including their knowledge composition separately in regressions unnecessarily complicates our estimation models, in the appendix, we show that our results and their interpretations are robust to the inclusion of the knowledge composition of these inventors as separate variables (Tables A5 and A6 in the appendix).

For each of the two top inventors in a team, we include two variables indicating their knowledge breadth and depth. These are based on technology classes of the patents where they are listed as inventors. Following past research (Boh, Evaristo and Ouderkirk, 2014; Fleming, Mingo and Chang, 2007), we use the number of unique technology classes in which an inventor had successfully filed patents during the five years prior to their focal patent as a proxy for the inventor's *knowledge breadth*. Here, the term "successfully filed" captures the application date of patents that were eventually granted. *Knowledge depth* is measured as the maximum number of patents the inventor had successfully filed in a single technology class during the same five-year period, a measure similar to that used by Boh, Evaristo and Ouderkirk (2014) and Mannucci and Yong (2018). The two measures together explain more than 90% of the variance in the inventor's 5-year patenting experience. Given that the mean and variance of inventors' knowledge depth and breadth varies across the four technologies, we de-mean all variables and normalize them based on the mean and standard deviation of each variable in each technology class.

We also measure the overlap between the knowledge of the two top inventors by calculating the ratio of technology classes in which they both successfully filed patents in the five years prior to the focal patent over relative to the number of unique technology classes in which either successfully filed patents (i.e., overlap over union).

To provide some intuition for how the measures are constructed, Figure 1 shows a graphical representation. In this example, both inventors have a knowledge breadth of four, meaning that they have successfully filed patents in four unique technology classes in the five years prior to their focal invention. Inventor 1's knowledge depth is five and inventor 2's knowledge depth is six. They also both have patents in technology classes D and C while the union of their patenting experience covers six technological classes, hence having an overlap ratio of one third.

Note that teams with the same level of aggregate patenting experience may differ substantially once the knowledge composition of each of their individual members is deconstructed. Figure 2 shows an example of two teams with similar aggregate levels of knowledge breadth and depth but different knowledge composition at the individual level. In team 1, the two inventors have distinct areas of knowledge with no overlap between them, whereas in team 2, the two inventors have similar and completely overlapping knowledge. When comparing the experience at the aggregated team level, the two inventors together have equal number of patents in the same technological classes across both teams. Thus, it is important to consider the individual experience of each inventor as well as the overlap between them.

-- Figures 1 and 2 about here --

Outcome Measures: Innovation Output

We focus on two types of the innovation output: whether the patent is an economic breakthrough and/or a novel breakthrough. One of the standard measures of breakthroughs in patent studies is based on future citations to the patent. While the measure is imperfect, several studies have shown strong correlation between the number of forward citations and the economic value of a patent (Trajtenberg, 1990; Hall et al., 2005). We use the number of citations received by a patent in a fixed window of time—here, five years since application date—as a proxy for its economic impact. Following past research (Hall et al., 2005; Ahuja and Lampert, 2001; Singh and Fleming 2010), we define economic breakthroughs as the top 10 percent most cited patents in the sample. We perform additional robustness tests using a count of the number of forward citations in five years after the application date as the dependent variable.

Because research has suggested that there are important differences between the citation-based measure of (economic) impact and the cognitive novelty of the patent (Kaplan and Vakili 2015), we include a measure of novel breakthroughs as a second measure of innovative outcomes. This is important because we might imagine that team configurations could have different relationships to cognitive novelty compared to the economic impact of a patent because these are produced through different processes. Cognitively novel patents are defined as those that introduce a new knowledge trajectory in a technological domain. The measure is based on Kuhn's (1962/1996) notion that shifts in ideas are reflected in shifts in language. Thus, the novelty in the vocabulary used to describe an idea can be used to assess the cognitive novelty of the idea itself.

Following Kaplan and Vakili's (2015) approach, we use topic modeling, an unsupervised automatic textual analysis method, to identify the set of topics present in the four sets of patents associated with the four selected technologies and identify which patents initiate each topic. Topic modeling is an increasingly used method in strategic management for studying large bodies of texts, particularly as associated with technologies (Croidieu & Kim, 2017; Wilson and Joseph, 2015; Hannigan et al 2019). Using the Stanford Topic Modeling Toolbox, we identified the 100 main topics represented by the abstracts of patents associated with each technology in our sample. Patent abstracts provide a summary of the novel aspects of an invention, are largely drafted by the inventors rather than patent lawyers, and are of approximately similar length, thus providing a useful basis of comparison across patents. The topic modeling algorithm produces a matrix that contains the association between each abstract and each of the 100 identified topics. Most patents contain only a few key topics, with the remaining topics having nearly no weight. Figure A1 in the appendix shows a sample abstract from our RFID sample and the top two topics.

Most topics are easy to recognize and interpret (see Tables A1 to A4 in the appendix for the top terms for the 100 topics identified in patents associated with MRI, RFID, stem cells, and nanotubes, respectively). For each technological domain, we also find a few topics that are not easily interpretable. For the sake of simplicity and reproducibility of results, we keep these topics in our sample. The inclusion of these topics can potentially increase the noise in our measure of cognitive novelty which would work against us finding significant effects. Once the topics are identified for each set of patents, following Kaplan and Vakili (2015), we select all patents over a 0.2 weighting threshold for each topic filed in the first 12 months since the first appearance of the topic. Using this approach, we identify 101, 153, 154, and 135 topic-originating patents in

MRI, RFID, stem cells, and nanotubes respectively. Based on this information, we construct the $TopicOriginating$ variable which is equal to one for topic-originating patents and 0 otherwise.

While we use both cognitive novelty (measure of topic origination) and economic impact (measure of top cited patents) as dependent variables in our estimations, it is important to note a gap between theoretical arguments concerning the role of knowledge recombination and measures of innovative output in prior studies. While these studies have largely estimated the impact of a certain type of knowledge recombination on economic impact (as proxied by forward citations), the mechanisms offered in these studies are largely concerned with how that type of knowledge recombination leads to novel inventions, which in turn, produces high economic impact (Fleming, 2001; Gittelman and Kogut, 2003). However, the relationship between knowledge recombination, novelty, and economic impact is not so straightforward. Scholars have documented many sources of tension created between pursuing novelty versus pursuing impact (Boudreau et al., 2016; Fleming, Mingo, and Chen, 2007). While novelty appears to be positively associated with economic impact, mechanisms that lead to novelty may not necessarily align with mechanisms that lead to economic impact. As Kaplan and Vakili (2015) show in the case of nanotechnology, recombination processes are positively associated with citations but negatively associated with cognitive novelty. Therefore, we should expect different relationships between team configuration and each of the two dependent variables we use.

Estimation Model

Because both dependent variables (novel breakthroughs and economic breakthroughs) are binary, we use a logistic model to estimate the association of team configuration with them. Our results are robust to using a linear model. In all estimations, we include a full set of interactions between technology and year fixed effects to control for the change in the opportunity landscape for each technology assuming that each technology might evolve at a different pace from the others. The interaction dummies ensure that each patent is compared to other patents in the same technology category and filed in the same year. In addition, we control for the number of claims on each patent and the number of references to prior patents as these have been shown to positively predict forward citation counts. For models with economic breakthrough as the dependent variable, we include the topic-originating patent indicator as an additional independent variable to examine whether novelty would be associated with economic impact.

In the first set of estimations, we interact the team configuration variables with each of the characteristics of technological domain—modularity and application breadth—separately. In the second set of regressions, we interact team configuration variables with indicators for each of the four technologies to understand how the combination of the two characteristics is associated with the relationship between the team configuration variables and the innovation outcomes.

For all estimations, we report the odds ratios. Several scholars have pointed out that the estimated interaction terms in non-linear models such as logits do not equal the marginal effects of interaction terms (Ai and Norton, 2003; Norton, Wang, and Ai 2004; Cornelißen and Sonderhof, 2009). However, here we report the multiplicative effects of interaction terms in terms of odds ratios. The multiplicative interpretations do not suffer from the issues raised by these scholars (Buis, 2010). Nevertheless, we have also replicated our results using a linear probability model as a robustness check.

RESULTS

Table 4 shows the summary statistics. Slightly fewer than 10% of patents are identified as economic breakthroughs (highly cited)[1] and approximately 3% of patents are novel breakthroughs (topic-originating). The average team size is 2.6, ranging from 2.4 to 3.1 across the four technological domains. In the five years prior to each focal patent, first inventors on average have filed patents in approximately 4.3 unique technology classes (the measure of knowledge breadth) and approximately 4.6 patents in the technology class in which they have the greatest number of patents (the measure of knowledge depth). First inventors in stem cells have significantly more knowledge breadth and depth than first inventors in other areas. Second inventors are substantially less experienced: they have on average filed patents in about 1.5 unique technology classes (breadth) and about 1.7 patents in the technology class in which they have the greatest number of patents (depth).

Note that in the regressions, we normalize measures of knowledge breadth and depth for the two top inventors within each technology class. The normalized measures have a mean of 0 and a standard deviation of 1. On average, the two top inventors have approximately 17% overlap in the set of technology classed in which they have filed patents. Third, fourth, and fifth inventors have on average 1.5, 0.6, and 0.3 patents filed in the five years prior to the focal patent. Patents in our sample list on average 20 claims and cite approximately 13 other patents as prior art. While there is significant variance in team configuration across technology classes, the variance within each technology class is greater than the variance across them.

-- Table 4 about here --

We start by comparing the independent association between team configuration and outcomes for broad vs. narrow application (Table 5), then modular vs. non-modular (Table 6) before showing the combined effects in Table 7. Table 5 presents the estimated relationship

[1] Due to the discrete nature of citation measures, the percentage of patents in the top 10% of citations for each technology is smaller than 10% overall.

between team configurations and each of the dependent variables in the case of a broadly-applied (RFID and nanotubes) or narrowly-applied (MRI or stem cells) technological domains. The results in the first column indicate that—other than a negative association between knowledge depth for narrowly-applied technologies—there is little association between the first inventor's knowledge breadth or depth and the degree to which broadly or narrowly-applied technologies are novel breakthroughs. However, the effects for the second inventor are intriguing. For broadly-applied technologies, we see a strong positive association between the knowledge breadth of the second inventor and the likelihood that the team's invention is cognitively novel (becomes a topic-originating patent). A standard deviation increase in the second inventor's knowledge breadth is associated with approximately 1.5 times increase in the chance of the team producing a topic-originating patent in a broadly-applied technological domain. In contrast, in narrowly-applied technologies, the second inventor's knowledge depth rather than the breadth is more important when producing cognitively novel patents. A standard deviation increase in the second inventor's knowledge depth is associated with approximately 1.4 times increase in the likelihood of producing a topic-originating patent in a narrowly-applied technological domain.

-- Table 5 about here --

The results are consistent with the idea that, for broadly-applied technologies, producing cognitively novel patents often involves seeking new application domains for the technology. The results suggest that second inventors—i.e., the inventors with the second most amount of experience on each patent—are more likely to be the ones who act as the bridge between the technology and a new application domain. In contrast, in narrowly-applied technologies, the second inventor's knowledge depth is more crucial to identify novel technological paths. Here, the novelty is generally associated with finding new methods of producing the technology or modifying the technology itself to be capable of new functionalities. Knowledge depth here can provide insight into the fundamentals of the technology and its attributes, potentially increasing the chance of finding novel breakthroughs. It is therefore not surprising that, in either case, the estimates for the overlap between the first and second inventors are not significant. It is the second inventor that provides a bridge to new domains when needed.

Turning to the economic impact in column 2, the results suggest that topic-originating patents are 1.4 times more likely to be economic breakthroughs compared with non-topic-originating patents. Moreover, the chances of producing an economic breakthrough grows with an increase in the first inventor's knowledge breadth in both broadly- and narrowly-applied technologies. The estimated effect is above and beyond the indirect effect of the first inventor's knowledge breadth on economic impact mediated through producing topic-originating patents. The effect is significantly larger for broadly-applied technologies. Very little else in team configuration matters for creating economic breakthroughs beyond the indirect effects mediated through producing topic-originating patents. Note that dropping the topic-originating patent variable from the regression in column 2 has a minimal effect on the team configuration variables.

One interpretation is that first inventors with wider knowledge breadth can diffuse the patented invention among a wider and more diverse audience. In other words, our measure of the inventor's knowledge breadth is simply a proxy for the breadth of the inventor's audience and reach. To the extent that this interpretation is accurate, we should expect that the first inventor's knowledge breadth would have a positive relationship with the chance of producing economic breakthroughs across any technological domain including all four we study here. We show this idea holds when we present the technological domain-level results in Table 7.

Table 6 presents the results for technological modularity: comparing modular technologies (RFID and MRI) with non-modular technological domains (stem cells and nanotubes). The estimates in column 1 suggest an interesting difference in the role of knowledge breadth for modular versus non-modular technological domains. In modular domains, teams producing cognitively novel patents seem to benefit from having first inventors with knowledge depth, and little overlap between the first inventor and the second inventor is needed. Here, the standardized protocols that exist in modular technologies can facilitate the knowledge integration between the team members. One might infer that too much knowledge overlap would lead to undesirable knowledge redundancy and reduce the chance of producing novel breakthroughs.

-- Table 6 about here --

However, in non-modular technologies, teams with a first inventor who has a greater knowledge breadth and substantial overlap in knowledge with the second inventor are significantly associated with producing cognitively novel breakthroughs. The estimates are consistent with the idea that in non-modular domains, the lead inventor acts as the knowledge integrator. Hence, it is important for the first inventor to cover a wider knowledge scope and have more overlap in knowledge with other inventors to be able to facilitate communication and knowledge integration at the team level. Overall, the results suggest that technological modularity can substitute for the first inventor's role as knowledge integrator.

The estimates for the association of team configuration with economic breakthroughs in the second column are similar to those reported previously in Table 5. Again, an increase in a first inventor's knowledge breadth is associated with a higher chance of producing economic breakthroughs in both modular and non-modular technological domains. The estimated effect is larger in non-modular technologies.[2]

Finally, in Table 7 we test how the combination of the two technological characteristics moderate the relationship between each team configuration and each innovation outcome. Each three-way interaction represents the effect of a particular team configuration variable in each of the four technological domains we study. That is, the four interactions between modularity and broad application variables map to the four technological domains in our sample. For example, the interaction term where the modularity dummy equals one and broad application dummy equals one represents the RFID technological domain. Hence, the three-way interactions between modularity, breadth of application, and each of the team configuration variables where modularity and broad application dummies both equal one, for example, essentially identifies the type of team configuration that is more likely to produce novel or economic breakthroughs in the RFID domain. Similarly, the three-way interactions between modularity and broad application dummies both equal to zero and team configuration variables indicate the type of team behind each type of breakthrough in the stem cell domain.

Column 1 shows the effects on novel breakthroughs (topic-originating patents). We discuss the findings for each technological domain separately. In the case of the MRI technologies, a narrow application and modular domain, the inventors' knowledge breadth has little effect on the chance of producing topic-originating patents. Given the narrow application domain of the technology, there is little that inventors can gain from having a greater breadth of knowledge. Meanwhile, since the technology is modular, standardized protocols can facilitate

[2] In Tables A5 and A6 in the appendix, we replicate the estimations in Tables 5 and 6, replacing the controls for the total patenting experience of the third, fourth and fifth inventors with their knowledge breadth and depth interacted with each technology characteristic (modularity and application breadth). The estimations reported here remain largely the same. Since there are few teams with more than two inventors that have produced novel or economic breakthroughs, the estimations for the third, fourth and fifth inventors are not reliable and may not represent the equilibrium association between team configuration and innovation outcomes.

knowledge integration at the team level, reducing the need for knowledge overlap between inventors' knowledge scope. Indeed, the estimates suggest that teams with lower levels of knowledge overlap are significantly more associated with producing topic-originating patents.

-- Table 7 about here --

Switching to RFIDs, a broadly-applied and modular technological domain, the estimated effect of the second inventor's knowledge breadth on the likelihood of producing topic-originating patents becomes positive and significant. A standard deviation increase in second inventor's knowledge breadth is associated with 1.4 times increase the chance of producing novel breakthroughs. At the same time, since the technology is modular, standardized protocols can facilitate knowledge integration at the team level. The estimated effect of knowledge overlap is negative and significant. That said, the negative effect of knowledge overlap is not as large as the effect we find for MRI patents. In the case of RFID, because of the large breadth of applications, inventors may find themselves in new application domains where the standardized interfaces are not adequate. In such situations, knowledge overlap may play an important role in knowledge integration at the team level.

In the case of stem cells, a narrowly-applied and non-modular technological domain, three effects stand out. Teams in which first inventor has more knowledge breadth are more associated with producing cognitively novel (topic originating) patents. The second inventor's knowledge depth is also significantly associated with the chance of producing topic-originating patents. Moreover, the effect of overlap between the first and second inventors' knowledge scope on the likelihood of producing cognitively novel patents is positive, large, and significant. Overall, the estimates suggest that in the absence of a modular design with standardized interfaces, an experienced inventor with wider knowledge breadth that overlaps with other team members is essential for knowledge integration at the team level. Meanwhile, due to the narrow-purpose range of integration, the team does not benefit much from the second inventor's knowledge breadth. Instead, an inventor with more knowledge depth is more likely to help the team produce cognitively novel innovations.

The results for nanotubes highlight the importance of both knowledge breadth and knowledge overlap in a broadly-applied but non-modular technological domain. The estimates suggest that teams in which both top inventors have more knowledge breadth are more likely to produce novel breakthroughs. This is consistent with the idea that for broadly-applied technologies, broader knowledge helps teams navigate a larger knowledge landscape and find new application areas. Meanwhile, because the technology is not modular, it is important for the top inventors to have knowledge overlap to be able to successfully integrate their diverse knowledge stocks.

The results of these three sets of analyses for the production of novel breakthroughs is summarized in Table 8. It shows that different team configurations would be useful in the case of both modular vs. non-modular technologies and broadly-applied vs. narrowly-applied technologies. Interestingly, some of these effects are amplified when they are combined, and some are offsetting. For example, for nanotubes, the inventors' knowledge breadth is important to bridge to new application domains given that the technology is broadly-applied. At the same time, due to the non-modular nature of technology, knowledge overlap allows inventors to integrate their diverse knowledge backgrounds. In contrast, in RFID which is also a broadly-applied, there is less need for such knowledge overlap since standardization can substitute the need for overlap as a facilitator of knowledge integration at the team level.

In stem cells, again we see the importance of knowledge overlap in the absence of a modular technological design. However, since the technology is narrow-purpose, second inventors with deeper knowledge are more likely to contribute to the production of cognitively novel (topic-originating) patents. Meanwhile, a first inventor who has the necessary knowledge breadth and overlap with the second inventor can facilitate the knowledge integration at the team level. Finally, we see that in the modular MRI domain, knowledge overlap is again negatively associated with the chance of producing novel breakthroughs. Here, the second inventor's knowledge breadth is also not helpful due to the narrow-purpose nature of the technology. A Wald test comparison of estimated coefficients in Table 7 shows significant differences on the main five dimensions of team configuration across the four domains at the .05 level.

To summarize the findings on novel breakthroughs, we show first in Table 8 the different effects of the types of technologies (broad vs. narrow, modular vs. non-modular) and how, in combination the effects are either reinforcing or offsetting. Figure 3 graphically compares the average de-meaned team configuration for the first two authors behind cognitively novel breakthroughs to the average de-meaned team configuration behind non-novel patents to visualize the differences in team configurations associated with novel breakthroughs relative to the average de-meaned team configuration (that does not produce a novel breakthrough).

-- Table 8 and Figure 3 about here --

For economic breakthroughs, we find exactly what was anticipated by the earlier analyses: the first inventor's knowledge breadth is key for achieving future citations, and this holds across all types of technology. The mirror effect is that the depth of knowledge of the first inventor is largely negatively associated with economic breakthroughs, though the size of the effect is small. There are few other consistent results for team configuration across the different domains, either for the second inventor's knowledge or for knowledge overlap. Because economic breakthroughs are measured by the number of forward citations, this finding suggests that, regardless of technological domain, inventors with more breadth of experience are more likely to garner more future citations to their patents. An implication is that breadth of experience captures other social processes such as extensiveness of inventor networks that might contribute to the uptake of particular ideas.

We performed several additional analyses to test the robustness of our results. Given that our results are based on interactions in a nonlinear estimation models, we have also tested the sensitivity of our findings to using a linear probability model instead of a logistic model. Table A7 in the appendix shows the estimates based on the linear probability model. The results are consistent with those based on the logistics model reported in Table 7.

Furthermore, one may be concerned that the effects we have attributed to the breadth of application might instead be driven by the difference between technological use versus development. In other words, unlike stem cell and MRI patents that largely involve some development of these technologies, RFID and nanotube technologies may simply enter into the innovation process as pure inputs without any additional development to the technologies themselves. Therefore, the effect of application breadth may be confounded. While these two attributes can be positively correlated, they are theoretically distinct. To address this concern, we selected a random subsample of patents in the nanotube and RFID domains and coded them under “pure input/use” and “development.” The first category contains all of the patents that use these technologies as input into the process. The second category includes all patents that involve some development or modification of these technologies. Our coding suggests that approximately 85% of the nanotube patents involve some development or modification of the nanotube technology. Using the same coding procedure, we found very similar figures for the

MRI and stem cell domains. Approximately 84% of stem cell patents and 82% of MRI patents involve some development or modification of these technologies.

For the RFID patents, the figure is lower at 55%. However, the coding revealed that within the RFID domain, there is a clear distinction between patents that mention the RFID technology in their titles or abstract versus those that mention the technology only in the claims section. In the former category, more than 90% of the patents involve some development of the RFID technology. To make sure our results are not driven by the variance in use vs. development, we did an additional robustness check excluding the latter category of RFID patents – i.e., the ones mentioning the technology only in the claims section. The results are reported in Table A8 in the appendix and are in line with those reported in Table 7. The only effect that changes from significant to insignificant is for the knowledge overlap between the first and second inventors. This is potentially driven by our focus on a subsample of RFID patents that heavily draw on the RFID technology in new application domains. On the one hand, the technological modularity lowers the need for knowledge overlap to facilitate knowledge integration at the team level. On the other hand, applying the technology in new application domains where standardized interfaces are inadequate may create a need for some knowledge overlap for lowering the communication and collaboration frictions. These two forces pull the effect of knowledge overlap in different directions. Note that in this robustness check, more than 84% of the nanotube patents and more than 90% of the RFID patents involve some development or modification of these technologies. The numbers are similar to those for the MRI and stem cell domains. It is thus very unlikely that our results regarding the application breadth are driven by the variance in use vs. development.

Finally, we also repeated our regressions in Table 7 on the full sample of patents included the 128 and 115 patents from the RFID and nanotubes samples that do not conform to the idea that the input technological domain and output technological domain are different for wider-ranging technologies. The results are reported in Table A9 in the appendix and are in line with those reported in Table 7.

It is important to note that our results are not causal and simply reveal associations between team configurations in each technological domain and the chance of producing novel or economic breakthroughs. Therefore, it is possible that the results might be driven by selection rather than treatment. Teams with different configurations could sort into the pursuit of different goals and hence our estimations may fully or partly capture this selection process. There are two ways to think about this selection mechanism. First, it is possible that the selection mechanisms are shaped by the underlying causal mechanisms. For example, a certain team configuration is tasked with pursuing novel breakthroughs because that type of configuration has more chance to achieve that goal. In such a case, the selection and causal mechanisms overlap, yet the estimates can still inform us about the causal mechanisms in play, though the size of the estimates would be biased.

Second, it is possible that the reason that a certain team configuration is tasked with pursuing a certain goal might be driven by factors other than the causal mechanisms linking the two. For example, it might be that managers care a lot about producing novel breakthroughs and hence put their most diverse teams in terms of experience together for that goal. At the same time, they support these projects financially which could hypothetically be the main reason why these teams succeed in producing novel breakthroughs. Then, what we see as the positive link between experience diversity and producing novel breakthroughs is simply a spurious correlation driven by the omitted variable of management and financial support. We cannot rule out this possibility entirely because of our empirical design and therefore caution our readers against drawing a causal interpretation.

However, our empirical design has a certain feature that limits the range of these alternative explanations considerably. Whatever alternative explanation one can think of, it should be able to explain the variation in the relationship between team configuration and breakthrough outputs of interest across the four domains. To continue on the previous example on managerial and financial support as an omitted variable, one should also be able to explain why managers in one domain would put a team with a diverse experience together to pursue novelty and then support it financially while managers in another domain would put a team with less than average diversity in experience together to pursue cognitive novelty and then support it financially. For us, it is difficult to think of such omitted variables that explain the relationship between breadth and depth of experience and breakthrough outputs consistently across the four domains. Yet, we cannot rule out the possibility that such alternative explanations may exist.

DISCUSSION AND CONCLUSION

This paper addresses tensions observed in prior research on organizing for innovation by conducting a cross-domain analysis of team configurations. While research has recognized that innovation has increasingly become a collaborative effort (Wuchty et al, 2007), especially as knowledge domains advance and individuals become more specialized (Jones, 2009; Agrawal, Goldfarb, and Teodoridis, 2016), research conflicts in its findings about what types of team configurations might be the most useful for creating innovative outputs. Do teams need to include inventors with knowledge breadth (Fleming, 2001; Gittelman and Kogut, 2003; Hargadon and Sutton, 1997; Uzzi and Spiro, 2005) or knowledge depth (Kaplan and Vakili, 2015; Kuhn 1962/1996; Taylor & Greve 2006; Weisberg, 1999)? Do teams need overlapping knowledge in order to integrate insights from diverse areas (Dahlin et al, 2005; Dougherty, 1992) or not (Burt, 2004)?

We wondered if the conflicting conclusions in prior results might come from differences in the characteristics of technologies studied, as these prior arguments have generally been agnostic to the technological domain for innovation, either drawing from a wide set of technological areas while only controlling for technology (e.g., Singh and Fleming, 2010; Arts and Veugelers, 2015) or focusing on a single technological domain, which may not generalize to other contexts (Phene et al, 2006; Rosenkopf and Nerkar, 2001).

We take a step towards resolving these tensions by asking how characteristics of a technological domain moderate the relationship between team configuration and innovation outcomes. Using a theoretically-driven 2x2 sample of patents in four technological domains, we found that different team configurations are associated with different technological characteristics. Drawing on the major distinctions in types of technologies made in the literature, we focused on technologies that vary on whether they are broadly or narrowly-applied (Bresnahan and Trajtenberg 1995, Helpman and Trajtenberg, 1994) and whether they are modular or non-modular (Langlois, 2002; Brusoni and Prencipe, 2001).

We operationalize these distinctions by studying patenting in four different technologies—magnetic resonance imaging (MRI) (modular and narrowly-applied), radio frequency identification (RFID) (modular and broadly-applied), stem cells (non-modular and narrowly-applied), and nanotubes (non-modular and broadly-applied). By examining associations between team configuration and innovative outcomes in these four types of technologies, we shed light on the contingent organizational designs associated with innovation.

Using two different measures of innovative outcomes—patents that represent either novel breakthroughs or economic breakthroughs—we look at how different team configurations are associated with each of the four technologies.

Team configurations for novel breakthroughs

With regard to cognitive breakthroughs, we find that modularity is a substitute for knowledge overlap in integrating diverse insights, while the first inventor serves as the integrator when technologies are not modular. Comparing broadly and narrowly-applied technologies, we find that the second inventor has a crucial role to play: for broadly-applied technologies in providing wide knowledge breadth, presumably to seek out a wide range of applications, or for narrowly-applied technologies, in greater knowledge depth, presumably to seek out anomalies. We also find that the interaction of modularity and application breadth moderates the effect of team configuration on the chance of producing novel breakthroughs.

These findings have three implications for theory and empirical analysis concerning organizing for innovation. First, we may not be able to create a general theory for team configurations that works in all settings. Attending to the different technological domains would be essential not only theoretically but also empirically: simply controlling for technological classes may be inadequate because the average effects in a broad range of technologies might disguise opposing effects within specific technologies.

Second, breadth of knowledge and depth of knowledge serve different creative functions and the degree to which they are either reinforcing or offsetting depends on the nature of the technology. For example, for broadly-applied technologies, the second inventor's breadth of knowledge is crucial for identifying new applications and depth of knowledge may reduce her ability to find these insights. For narrowly-applied technologies, the second inventor's depth of knowledge is precisely what allows for the new insights in the narrow domain and breadth of knowledge may hinder the ability to seek out those anomalies.

Third, creativity depends on recombinations of insights; however, this integration process can be accomplished in different ways depending on the nature of the technology. In non-modular technologies, inventors must have the breadth of knowledge and substantial overlap in the team in order to integrate ideas. However, in modular technologies, standards can substitute for the human integration function. Further, these effects are sensitive to whether the technology is broadly or narrowly-applied.

We should emphasize that, given our empirical design, our estimations only show the association between team configuration and innovation output and may not necessarily be based on causal relationships. On the one hand, it is possible that our results do reveal a causal relationship between team configuration and each innovation type. On the other hand, our results might be driven by some selection mechanism—that inventors with certain specialization profiles might be more likely to collaborate for certain type of outcomes (novel or impactful). For example, it is possible that in the RFID domain, inventors with a deep knowledge seek out inventors with a wider knowledge breadth if they plan to work on more novel ideas. Or, alternatively, managers who put teams with such a profile together may also task them with pursuing novel innovations. Such selection mechanisms may themselves reflect the underlying causal effects: that because deep knowledge needs to be combined with knowledge breadth to produce novelty in the RFID domain, inventors or their managers try to select on such team profiles if their aim is to produce novelty. Alternatively, such selection mechanisms might be driven by other omitted factors unrelated to the underlying causal effects at work.

However, any such explanation would have to be consistent with the variance we see in the relationship between team configuration and innovation output across all four technological domains. In other words, such an omitted factor should, on the one hand, lead to the formation of teams with complementary breadth and depth between the first and second inventors in the RFID domain while at the same time increase the chance of such teams producing novel innovations, and, on the other hand, lead to the formation of teams with similar knowledge profiles (more depth than breadth) in the stem cell area while, again, at the same time increase the likelihood of producing novel innovations by such teams in that domain. While we cannot think of a plausible selection mechanism that would provide such an alternative explanation, given our empirical design, we cannot rule out all such selection mechanisms.

Moreover, given that we use granted patents as a measure of innovation, our results and interpretations are conditioned on the success of teams to produce patentable inventions. Nevertheless, our main conclusions based on the moderating role of the degree of modularity and the breadth of application still hold.

Team configurations for economic breakthroughs

We show that novel breakthroughs are positively associated with economic impact across all four technological domains. Much of the patent literature on creating breakthroughs has hypothesized but not measured this effect directly (e.g., Trajtenberg *et al* 1997; Singh & Fleming 2010; Phene *et al* 1997). We are able to show empirically that having more novel ideas is associated with the likelihood of creating an economic breakthrough no matter the domain (at least in the four we have studied).

When we turn to what team configurations contribute to these economic impacts, we find that it is primarily the breadth of experience of the first inventor, and this holds for all of the technological domains we studied. These results may suggest that the breadth of inventor experience may also be capturing something about the social processes of diffusion of ideas in which more highly connected inventors are more likely to get cited. The inventor's knowledge breadth may act as a proxy for the breadth of the inventor's audience and reach. The fact that, in our analyses, the first inventor's knowledge breadth has a positive relationship with the chance of producing economic breakthroughs across the four diverse technologies we study here offers some support for this conclusion. However, we do not have a strong theoretical underpinning to interpret these results pertaining to the relationship between knowledge composition of teams and the economic impact of their innovations. Since we have drawn our theoretical arguments from a literature that most directly speaks to the relationship between the knowledge composition of teams and the novelty of their innovations, our interpretations with regard to the economic impact of team configurations are much more speculative and call for future research to clarify the relationship between team configuration, novelty and economic impact.

In conclusion, we find that team configurations for innovation are contingent on the technological domain in which an organization operates. Though early continency theory (Woodward 1958, 1965) highlighted that organizational form should depend on the types of technologies an organization pursues, much of the existing literature on team configurations for innovation has not addressed these contingencies directly. Our cross-domain analysis demonstrates that abstracting away from the role of technological domain can lead to theoretical and empirical confusion, where breadth, depth and overlap of knowledge in teams is useful in some domains and not others. Our study offers insight into how those contingencies operate and resolves some of the inconclusive findings about the nature of the relationships between team design choices and innovative outcomes.

REFERENCES

Agrawal, A., Goldfarb, A., & Teodoridis, F. (2016). Understanding the changing structure of scientific Agrawal, A., Goldfarb, A., & Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics*, *8*(1), 100-128.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, *80*(1), 123-129.

Ahuja, G., & Lampert, C. M. (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic management journal*, *22*(6-7), 521-543.

Arts, S., & Veugelers, R. (2014). Technology familiarity, recombinant novelty, and breakthrough invention. *Industrial and Corporate Change*, *24*(6), 1215-1246.

Audia, P. G., & Goncalo, J. A. (2007). Past success and creativity over time: A study of inventors in the hard disk drive industry. *Management Science*, *53*(1), 1-15.

Bercovitz, J., & Feldman, M. (2011). The mechanisms of collaboration in inventive teams: Composition, social networks, and geography. *Research Policy*, *40*(1), 81-93.

Berman, S. L., Down, J., & Hill, C. W. (2002). Tacit knowledge as a source of competitive advantage in the National Basketball Association. *Academy of management Journal*, *45*(1), 13-31.

Bikard, M., Vakili, K., & Teodoridis, F. (2018). When collaboration bridges institutions: The impact of university–industry collaboration on academic productivity. *Organization Science,* 30(2), 426-445.

Boh, W. F., Evaristo, R., & Ouderkirk, A. (2014). Balancing breadth and depth of expertise for innovation: A 3M story. *Research Policy*, 43(2), 349-366.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. Management Science, 62(10), 2765-2783.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. *Journal of econometrics*, *65*(1), 83-108.

Brusoni, S., & Prencipe, A. (2001). Unpacking the black box of modularity: technologies, products and organizations. *Industrial and Corporate Change*, *10*(1), 179-205.

Brusoni, S., Prencipe, A., & Pavitt, K. (2001). Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make?. Administrative science quarterly, 46(4), 597-621.

Buis, M. L. (2010). Direct and indirect effects in a logit model. *The Stata Journal*, *10*(1), 11-29.

Burt, R. S. (2004). Structural holes and good ideas. *American journal of sociology*, *110*(2), 349-399.

Cornelißen, T., & Sonderhof, K. (2009). Partial effects in probit and logit models with a triple dummy-variable interaction term. *The Stata Journal*, *9*(4), 571-583.

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science*, *12*(3), 346-371.

Croidieu, G., & Kim, P. H. (2017). Labor of love: Amateurs and lay-expertise legitimation in the early US radio field. Administrative Science Quarterly, forthcoming.

Dahlin, K. B., Weingart, L. R., & Hinds, P. J. (2005). Team diversity and information use. *Academy of Management Journal*, *48*(6), 1107-1123.

Dougherty, D. (1992). Interpretive barriers to successful product innovation in large firms. *Organization science*, *3*(2), 179-202.

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

Fleming, L., Mingo, S., & Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. Administrative science quarterly, 52(3), 443-475.

Gittelman, M., & Kogut, B. (2003). Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science*, *49*(4), 366-382.

Grant, R. M. (1999). Prospering in dynamically-competitive environments: Organizational capability as knowledge integration. In Knowledge and strategy (pp. 133-153).

Hall, B. H., & Trajtenberg, M. (2004). Uncovering GPTs with patent data. NBER Working Paper w10901.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. RAND Journal of economics, 16-38.

Hannigan, Tim, Richard F. J. Haans, Keyvan Vakili, Hovig Tchalian, Vern L. Glaser, Milo Wang, Sarah Kaplan, and P. Devereaux Jennings (2019), Topic Models in Management Research: Rendering New Theory from Big Textual Data, *Academy of Management Annals*, 13(2) 586-632.

Hargadon, A., & Sutton, R. I. (1997). Technology brokering and innovation in a product development firm. *Administrative science quarterly*, 716-749.

Helpman, E., & Trajtenberg, M. (1994). *A time to sow and a time to reap: Growth based on general purpose technologies*. NBER Working Paper w4854.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization science*, *21*(5), 1016-1033.

Jones, B. F. (2009). The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder?. *The Review of Economic Studies*, *76*(1), 283-317.

Kaplan, S., & Vakili, K. (2015). The double‐edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435-1457.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative science quarterly*, 81-104.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, *3*(3), 383-397.

Krauss, R. M., & Fussell, S. R. (1990). Mutual knowledge and communicative effectiveness. *Intellectual teamwork: Social and technological foundations of cooperative work*, 111-146.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lai, R., D'Amour, A., Yu, A. Sun, Y., Fleming, L. 2011. "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)", hdl:1902.1/15705, Harvard Dataverse, V5.

Langlois, R. N. (2002). Modularity in technology and organization. *Journal of economic behavior & organization*, *49*(1), 19-37.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335-1342.

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American sociological review*, *72*(4), 533-561.

Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, *62*(1), 105-139.

Leiponen, A., & Helfat, C. E. (2011). Location, decentralization, and knowledge sources for innovation. *Organization Science*, *22*(3), 641-658.

Mannucci, P. V., & Yong, K. (2018). The differential impact of knowledge depth and knowledge breadth on creativity over individual careers. *Academy of Management Journal*, *61*(5), 1741-1763.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, *2*(1), 71-87.

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, *4*(2), 154-167.

Østergaard, C. R., Timmermans, B., & Kristinsson, K. (2011). Does a different view create something new? The effect of employee diversity on innovation. *Research Policy*, *40*(3), 500-509.

Perrow, C. (1967). A framework for the comparative analysis of organizations. *American sociological review*, 32(2), 194-208.

Perry-Smith, J. E., & Shalley, C. E. (2014). A social composition view of team creativity: The role of member nationality-heterogeneous ties outside of the team. *Organization Science*, *25*(5), 1434-1452.

Phene, A., Fladmoe-Lindquist, K., & Marsh, L. (2006). Breakthrough innovations in the US biotechnology industry: the effects of technological space and geographic origin. *Strategic Management Journal*, *27*(4), 369-388.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative science quarterly*, 116-145.

Shea, C. M., Grinde, R., & Elmslie, B. (2011). Nanotechnology as general-purpose technology: empirical evidence and implications. Technology Analysis & Strategic Management, 23(2), 175-192.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality?. *Management science*, *56*(1), 41-56.

Rosenkopf, L., & Nerkar, A. (2001). Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal*, *22*(4), 287-306.

Sampson, R. C. (2007). R&D alliances and firm performance: The impact of technological diversity and alliance organization on innovation. *Academy of management journal*, *50*(2), 364-386.

Sanchez, R., & Mahoney, J. T. (1996). Modularity, flexibility, and knowledge management in product and organization design. *Strategic management journal*, *17*(S2), 63-76.

Schilling, M. A., & Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, *40*(10), 1321-1331.

Smith, K. G., Collins, C. J., & Clark, K. D. (2005). Existing knowledge, knowledge creation capability, and the rate of new product introduction in high-technology firms. *Academy of management Journal*, *48*(2), 346-357.

Stan, M., & Puranam, P. (2017). Organizational adaptation to interdependence shifts: The role of integrator structures. Strategic Management Journal, 38(5), 1041-1061.

Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, *112*(4), 639.

Taylor, A., & Greve, H. R. (2006). Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, *49*(4), 723-740.

Teodoridis, F., Bikard, M., & Vakili, K. (2018). Creativity at the Knowledge Frontier: The Impact of Specialization in Fast- and Slow-paced Domains. *Administrative Science Quarterly*. https://doi.org/10.1177/0001839218793384

Thompson, J. D. (1967). *Organizations in Action*. New York: McGraw Hill.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. The Rand Journal of Economics, 172-187.

Ulrich, K. (1995). The role of product architecture in the manufacturing firm. *Research policy*, *24*(3), 419-440.

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American journal of sociology*, *111*(2), 447-504.

Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed.), Handbook of creativity: 226-250. New York: Cambridge University Press

Wilson, A. J., & Joseph, J. (2015). Organizational attention and technological search in the multibusiness firm: Motorola from 1974 to 1997. In Cognition and Strategy (pp. 407-435). Emerald Group Publishing Limited.

Woodward, J. (1958), *Management and Technology*, London: H.M.S.O

Woodward, J. (1970). *Industrial organization: Behaviour and control*. Oxford University Press

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036-1039.

Figure1 1: Team knowledge composition dimensions

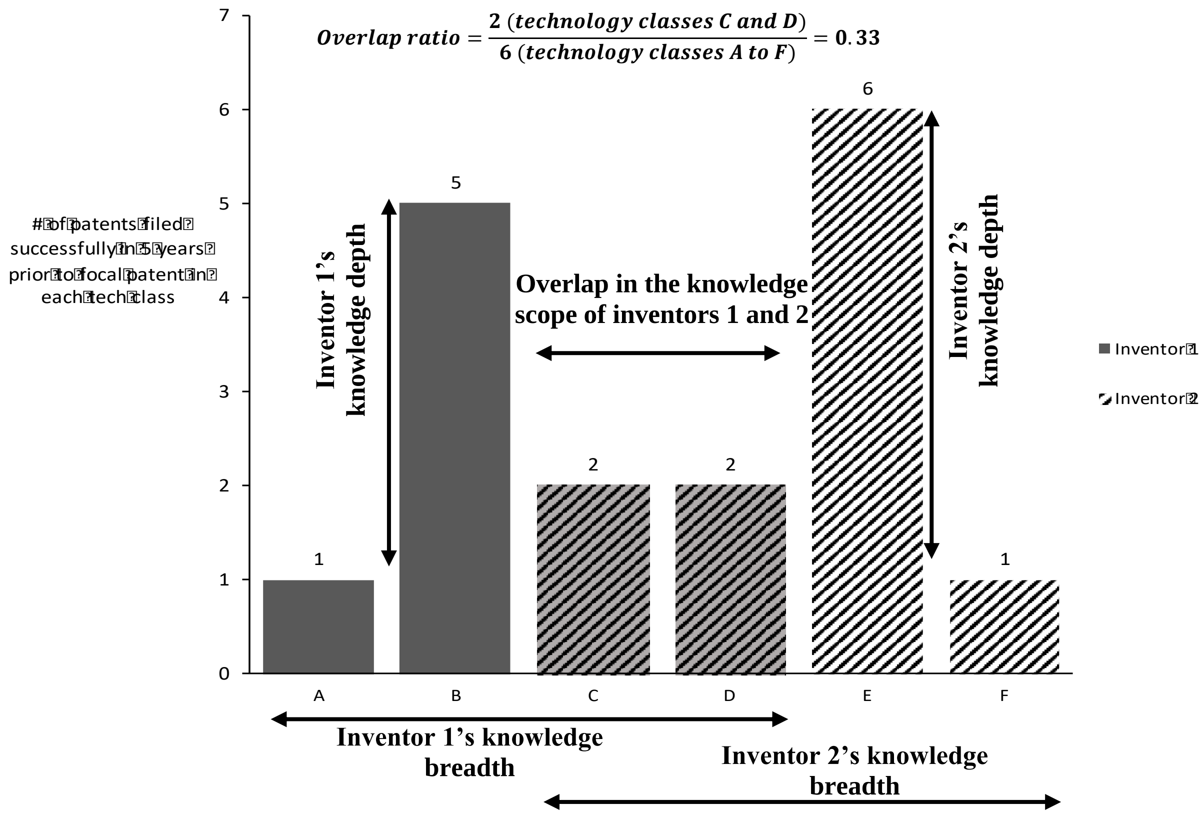


Figure 2: Sample teams with same aggregated patenting experience but different knowledge compositions

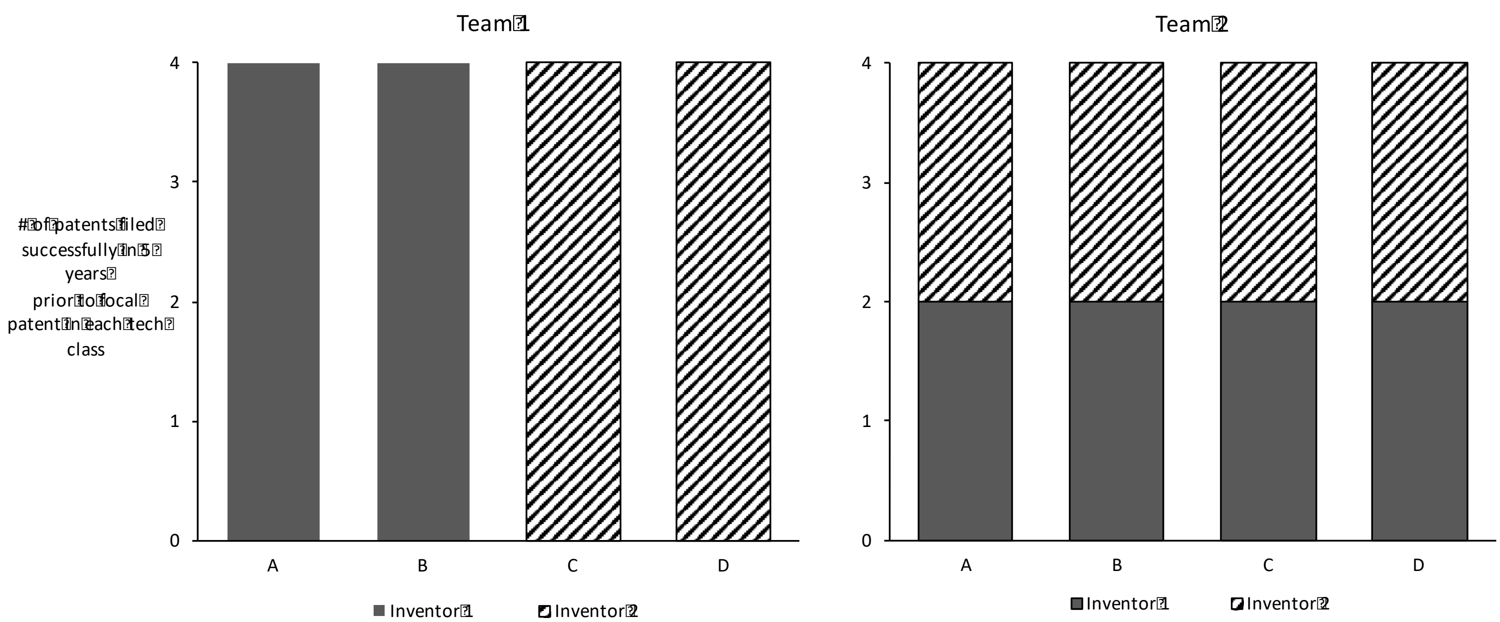
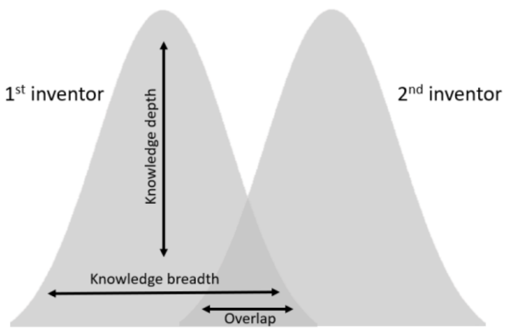
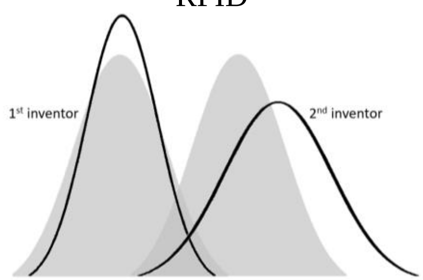


Figure 3: Illustrating the results graphically: 1st and 2nd inventor knowledge breadth, depth and overlap as associated with novel breakthroughs

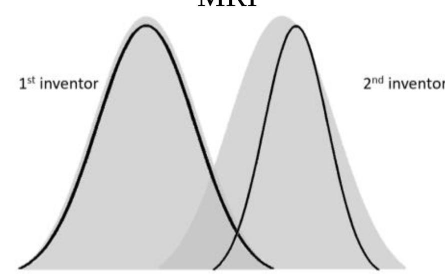


This figure shows a sample de-meaned schematic of the first and second inventors' knowledge profiles in an average team that has not produced a novel breakthrough in our sample. Note that the 1st inventor has, on average, much more experience than the 2nd inventor across all four domains in our sample. However, because we do not compare 1st inventors to 2nd inventors in our analysis, we have de-meaned both to the same level (with a mean of 1 and a standard deviation of 1). The height of each curve shows the knowledge depth and the width shows the knowledge breadth of each inventor. The overlapping area between the two inventors indicates the amount of overlap in their knowledge. In the graphs below, we overlay the de-meaned profile of the 1st and 2nd inventors on novel patents in each domain on this baseline figure.

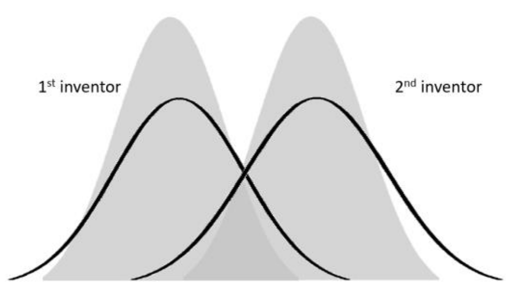
RFID



MRI



Nanotubes



Stem cell

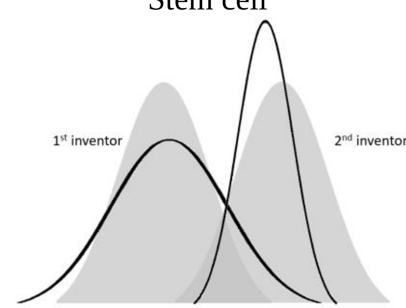


Table 1: Characteristics of the selected four technologies

	Broadly-applied	Narrowly-applied
Modular	RFID	MRI
Non-Modular	Nanotube	Stem Cell

Table 2: Search terms and results for each technology

	SEARCH TERMS	NUMBER OF RESULTS FROM EACH SEARCH	TOTAL NUMBER OF RESULTS FROM ALL SEARCHES	NUMBER OF RESULTS AFTER MERGING PATENTS IN THE SAME PATENT FAMILY
MRI	Terms searched in abstract, title, and claims on USPTO: MRI; fMRI; nMRI; "MR imaging;" "nMR imaging;" "magnetic resonance"	10,337	10,337	9,230
RFID	Terms searched in abstract, title, and claims on USPTO: RFID; "radio frequency identification"	4,009	4,059	3,521
	All patents in technological classes designated to RFID patents: 340/13.26	128		
Stem cell	Terms searched in abstract, title, and claims on USPTO: "stem cell;" hESC; iPSC	1309	2366	1,775
	All patents listed on the StemCellPatents.com, granted before 2010	503		
Nanotube	Terms searched in abstract, title, and claims on USPTO: nanotube; fullerene	1585	2,826	2,384
	All patents in USPTO cross-reference classes designated to nanotubes: 977/735-752	305		
	All patents in Derwent technological classes designated to nanotubes: B05-U, C05- U, E05-U, E31-U02, L02-H04B, U21-C01T, X12-D02C2D, X12-D07E2A, X12-E03D, X16-E06A1A.	1057		

Table 3: Average application breadth and its growth over time for each technology

	Average number of unique citing 3-digit technological classes	Growth in the number of unique citing 3-digit technological classes
MRI (narrowly-applied)	3.4	5%
RFID (broadly-applied)	6.2	24%
Stem cell (narrowly-applied)	3.9	0%
Nanotube (broadly-applied)	6.8	18%

Table 4: Descriptive features of four technologies

	All patents	MRI	RFID	Stem cells	Nanotubes
Economic breakthroughs (top 10% cited)	0.079 (0.270)	0.098 (0.297)	0.096 (0.295)	0.086 (0.281)	0.094 (0.292)
Novel breakthroughs (topic-originating patents)	0.026 (0.161)	0.011 (0.104)	0.043 (0.204)	0.087 (0.282)	0.057 (0.231)
Number of inventors	2.623 (1.751)	2.443 (1.618)	2.517 (1.795)	2.991 (1.955)	3.082 (1.991)
First inventor's knowledge breadth (non-normalized)	4.284 (5.949)	3.504 (4.602)	4.219 (6.773)	3.325 (4.517)	7.424 (8.562)
First inventor's knowledge depth (non-normalized)	4.603 (13.450)	4.247 (6.553)	4.120 13.384	3.422 (5.960)	7.688 (31.333)
Second inventor's knowledge breadth (non-normalized)	1.508 (2.900)	1.219 (2.360)	1.437 (3.520)	1.172 (2.119)	2.816 (4.121)
Second inventor's knowledge depth (non-normalized)	1.709 (4.951)	1.591 (4.165)	1.337 (5.094)	1.237 (3.259)	2.883 (8.335)
Overlap in the knowledge scope of the first and second inventors	0.167 (0.290)	0.175 (0.303)	0.101 (0.228)	0.180 (0.314)	0.198 (0.288)
Number of claims	19.860 (16.709)	18.935 (16.024)	22.291 (15.775)	19.977 (19.528)	21.237 (19.250)
Number of backward references to patents	13.366 (33.642)	11.008 (25.414)	23.505 (54.928)	11.822 (29.993)	15.999 (37.325)
Third inventor's 5-year patenting experience	1.504 (6.713)	1.183 (5.181)	1.438 (8.007)	1.181 (5.248)	3.131 (11.271)
Fourth inventor's 5-year patenting experience	0.601 (4.100)	0.462 (3.057)	0.406 (2.959)	0.436 (2.938)	1.523 (7.435)
Fifth inventor's 5-year patenting experience	0.255	0.202	0.165	0.213	0.743

	(2.498)	(1.821)	(1.853)	(2.382)	(5.359)

Table 5: Impact of team configurations on innovation outcomes: broadly and narrowly-applied technologies

Estimation Model:	**Logistic (Odds Ratios Reported)**	
DV:	**Novel breakthrough (1)**	**Economic breakthrough (2)**
Novel breakthroughs		1.356
		(0.014)
First inventor's knowledge breadth ×		
Narrowly-applied	1.289	1.050
	(P=0.102)	(P=0.087)
Broadly-applied	1.022	1.211
	(P=0.892)	(P=0.000)
First inventor's knowledge depth ×		
Narrowly-applied	0.841	1.054
	(P=0.038)	(P=0.164)
Broadly-applied	0.815	0.959
	(P=0.656)	(P=0.150)
Second inventor's knowledge breadth ×		
Narrowly-applied	0.724	0.984
	(P=0.024)	(P=0.746)
Broadly-applied	1.470	0.986
	(P=0.000)	(P=0.607)
Second inventor's knowledge depth ×		
Narrowly-applied	1.364	0.992
	(P=0.005)	(P=0.821)
Broadly-applied	0.586	0.982
	(P=0.000)	(P=0.784)
Overlap in the knowledge scope of the first and second inventors ×		
Narrowly-applied	0.720	1.038
	(P=0.728)	(P=0.713)
Broadly-applied	1.008	0.909
	(P=0.941)	(P=0.337)
Full set of controls	Yes	Yes
Technology-time fixed effects	Yes	Yes
Number of observations	4,001	16,744
R-squared	0.269	0.182

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table 6: Impact of team configurations on innovation outcomes: modular and non-modular technologies

	Estimation Model:	**Logistic (Odds Ratios Reported)**	
	DV:	**Novel breakthrough (1)**	**Economic breakthrough (2)**
Novel breakthroughs			1.322 (0.033)
First inventor's knowledge breadth ×			
	Non-Modular	1.369 (P=0.007)	1.239 (P=0.000)
	Modular	0.908 (P=0.002)	1.045 (P=0.039)
First inventor's knowledge depth ×			
	Non-Modular	0.722 (P=0.000)	0.855 (P=0.000)
	Modular	0.986 (P=0.863)	1.066 (P=0.002)
Second inventor's knowledge breadth ×			
	Non-Modular	0.802 (P=0.464)	1.028 (P=0.507)
	Modular	1.294 (P=0.013)	0.978 (P=0.322)
Second inventor's knowledge depth ×			
	Non-Modular	1.318 (P=0.040)	1.149 (P=0.001)
	Modular	0.786 (P=0.069)	0.974 (P=0.584)
Overlap in the knowledge scope of the first and second inventors ×			
	Non-Modular	1.612 (P=0.069)	1.131 (P=0.368)
	Modular	0.248 (P=0.064)	0.965 (P=0.197)
Full set of controls		Yes	Yes
Technology-time fixed effects		Yes	Yes
Number of observations		4,001	16,744
R-squared		0.271	0.183

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table 7: Impact of team configurations on innovation outcomes: all four technologies

Estimation Model:	**Logistic (Odds Ratios Reported)**	
DV:	**Novel breakthrough (1)**	**Economic breakthrough (2)**
Novel breakthroughs		1.332 (0.027)
First inventor's knowledge breadth ×		
modular × narrowly-applied (MRI)	1.002 (P=0.704)	1.032 (P=0.000)
modular × broadly-applied (RFID)	0.853 (P=0.000)	1.139 (P=0.000)
non-modular × narrowly-applied (Stem Cells)	1.556 (P=0.000)	1.162 (P=0.000)
non-modular × broadly-applied (Nanotubes)	1.413 (P=0.000)	1.292 (P=0.000)
First inventor's knowledge depth ×		
modular × narrowly-applied (MRI)	0.932 (P=0.000)	1.080 (P=0.000)
modular × broadly-applied (RFID)	1.224 (P=0.000)	0.996 (P=0.641)
non-modular × narrowly-applied (Stem Cells)	0.745 (P=0.000)	0.880 (P=0.000)
non-modular × broadly-applied (Nanotubes)	0.189 (P=0.000)	0.884 (P=0.000)
Second inventor's knowledge breadth ×		
modular × narrowly-applied (MRI)	0.788 (P=0.000)	0.974 (P=0.938)
modular × broadly-applied (RFID)	1.383 (P=0.000)	0.981 (P=0.441)
non-modular × narrowly-applied (Stem Cells)	0.600 (P=0.000)	1.094 (P=0.058)
non-modular × broadly-applied (Nanotubes)	1.531 (P=0.000)	1.010 (P=0.410)
Second inventor's knowledge depth ×		
modular × narrowly-applied (MRI)	0.949 (P=0.375)	1.002 (P=0.938)
modular × broadly-applied (RFID)	0.680 (P=0.000)	0.914 (P=0.019)
non-modular × narrowly-applied (Stem Cells)	1.536 (P=0.000)	1.002 (P=0.948)
non-modular × broadly-applied (Nanotubes)	0.378 (P=0.000)	1.200 (P=0.000)
Overlap in the knowledge scope of the first and second inventors ×		
modular × narrowly-applied (MRI)	0.158 (P=0.000)	0.973 (P=0.506)
modular × broadly-applied (RFID)	0.898 (P=0.014)	0.762 (P=0.011)
non-modular × narrowly-applied (Stem Cells)	2.289 (P=0.000)	1.360 (P=0.000)
non-modular × broadly-applied (Nanotubes)	1.164 (P=0.000)	0.994 (P=0.849)
Full set of controls & Technology-time fixed effects	Yes	Yes
Number of observations	4,001	16,744
Pseudo R-squared	0.279	0.184

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table 8: Team configuration for novel breakthroughs

		Broadly-applied	Narrowly-applied
		• 2^{nd} inventor acts as a bridge to new applications	• 2^{nd} inventor's knowledge depth helps identify novelty
Modular	• Standard interfaces between components substitute for knowledge overlap in facilitating team-level knowledge integration	**RFID** • 2^{nd} inventor acts as a bridge to new applications • Little knowledge overlap required to integrate	**MRI** • 2^{nd} inventor's knowledge depth helps identify novelty • Little knowledge overlap required to integrate
Non-Modular	• Knowledge overlap is essential to coordinate across inventors • 1^{st} inventor's knowledge breadth is needed for team-level knowledge integration	**Nanotube** • All inventors need knowledge breadth to identify new applications and facilitate team-level knowledge integration • Some degree of knowledge overlap required to coordinate across inventors	**Stem Cells** • 1^{st} inventor's knowledge breadth is in facilitating team-level knowledge integration • 2^{nd} inventor's knowledge depth helps identify novelty • Knowledge overlap is essential to coordinate across inventors

ONLINE APPENDIX
ORGANIZING FOR INNOVATION:
HOW TEAM CONFIGURATIONS VARY WITH MODULARITY AND BREADTH OF APPLICATION

Figure A1: Sample RFID patent abstract with associated top three identified topics

Patent Number: US 7540413
Title: Radio frequency identifiers in game tickets
Inventors: Meehan; Richard; Carney; Stephen; Seymour; Jennifer; Fitzgerald; Craig; Finocchio; Richard
Assignee: Gtech Rhode Island Corporation (Providence, RI)
Application Date: November 24, 2003, **Issue Date**: June 2, 2009
USPTO classifications: **235/381**; 283/903; 463/17; 902/23
Abstract: Systems and methods of distributing, dispensing and validating game tickets provide for the use of radio frequency identifiers (RFIDS). Game ticket consumables may also be tracked with RFIDS. Ticket data such as game numbers, void if removed numbers, theme descriptions, place styles, price points and player account information can be stored to a memory of a game ticket, where the ticket is capable of transmitting the ticket data as an RF signal. Approaches also provide for the use of RFIDS in game sponsor/ticket printer facilities, ticket warehouse facilities and ticket destruction facilities.

Identified topics
Topic 70 (with a topic weight of 95%):
Top terms: gaming, game, play, table, calibration, player, clip, participants, ride, players

Topic 7 (with a topic weight of 5%):
Top terms: memory, stored, storing, cell, operation, non-volatile, write, coded, writing, radio-frequency
The rest: approximately 0% each

Table A1- Top 10 terms of the identified 100 topics in abstracts of MRI patents

Topic 1	reference	calibration	movement	displacement	automatically	phantom	tracking	comparing	optimal	comparison
Topic 2	transfer	biological	acid	nucleic	amino	sequence	acids	sequences	probes	gene
Topic 3	amplitude	high-frequency	constant	b.sub.1	b.sub.0	oscillation	frequency	oscillator	oscillating	frequencies
Topic 4	examination	plane	planes	imaged	whole-body	tomography	cross-sectional	perpendicular	images	lying
Topic 5	flow	blood	velocity	flowing	vessels	perfusion	arterial	stationary	vessel	angiography
Topic 6	heart	cardiac	blood	artery	coronary	patient	plaque	wall	myocardial	ventricle
Topic 7	instrument	procedure	surgical	coordinate	medical	patient	interventional	procedures	tracking	navigation
Topic 8	magnetization	pulses	spins	angle	180.degree	90.degree	flip	sequence	slab	transverse
Topic 9	images	x-ray	organ	computed	registration	medical	tomographic	diagnostic	tomography	registered
Topic 10	core	monitoring	mri	environment	sensors	article	devices	monitor	battery	techniques
Topic 11	scan	fourier	projection	three	dimensional	transformation	transform	two-dimensional	domain	recording
Topic 12	probe	catheter	tip	distal	balloon	guide	elongated	lumen	probes	shaft
Topic 13	values	pixel	pixels	voxel	calculated	voxels	vector	calculating	raw	vectors
Topic 14	magnet	assembly	pole	magnets	gap	yoke	assemblies	pair	ferromagnetic	uniform
Topic 15	echo	spin	echoes	sequence	phase-encoding	read	phase	readout	train	pulses
Topic 16	materials	human	injection	hybrid	tissues	carbon	anatomic	shear	mixing	implant
Topic 17	measurement	measured	distribution	correction	corrected	distortion	measurements	factor	correcting	calculated
Topic 18	diffusion	weighted	species	waveform	spectral	tensor	images	weighting	waveforms	alpha
Topic 19	phase	encoding	error	encode	errors	phases	navigator	correction	encoded	readout
Topic 20	polymer	molecular	animal	chain	acid	gel	chemical	structure	structural	represented
Topic 21	digital	filter	noise	filtered	converter	analog	filtering	frequency	pass	low
Topic 22	nuclear	nuclei	spin	free	induction	decay	spins	excited	atomic	electron
Topic 23	array	detector	local	module	tomography	emission	pet	modules	phased	positron
Topic 24	gas	cavity	chamber	valve	hyperpolarized	stent	pressure	toroid	vapor	noble
Topic 25	shim	negative	shimming	homogeneity	inhomogeneity	positive	strength	magnet	shims	inhomogeneities
Topic 26	fluid	bone	fluids	media	property	porous	reservoir	measurements	rock	pore
Topic 27	output	input	amplifier	port	stage	outputs	circuit	ports	feedback	voltage
Topic 28	mechanism	holder	hole	compression	equipment	subsystem	integrated	motor	isolation	driving
Topic 29	resin	weight	oil	resistance	copolymer	polymer	film	ppm	temperature	ethylene
Topic 30	reaction	organic	catalyst	acid	salt	solution	rare	compound	mixture	fraction
Topic 31	coils	pair	pairs	transmit	quadrature	receive	array	three	currents	saddle
Topic 32	member	hollow	disc	alignment	spinal	movable	cover	positioning	appliance	interior
Topic 33	container	external	flexible	mechanical	vibration	vibrations	noise	air	force	connecting
Topic 34	frequency	radio	frequencies	band	resonant	larmor	receiving	tuned	center	dipoles
Topic 35	cells	therapeutic	antibody	cancer	compounds	targeting	binding	agent	diagnostic	treatment
Topic 36	water	fat	t.sub.1	saturation	suppression	proton	t.sub.2	quantum	images	states
Topic 37	magnet	superconducting	flux	path	dipole	partial	operation	coils	radially	shielding

Topic 38	formation	tool	measurements	logging	borehole	earth	nmr	formations	nuclear	porosity
Topic 39	source	light	beam	radiation	sources	laser	optical	mirror	detector	beams
Topic 40	product	shape	disease	variables	segmented	status	patients	comparison	hand	age
Topic 41	axial	winding	cylindrical	windings	length	wound	radial	z-axis	gradients	bobbin
Topic 42	sequence	slice	pulses	slices	sequences	inversion	saturation	images	recovery	after
Topic 43	brain	living	stimulation	functional	window	variation	physiological	stimulus	indicator	neural
Topic 44	spectrum	chemical	shift	peak	spectra	spectroscopy	nuclear	spectrometer	spectroscopic	peaks
Topic 45	technique	gradients	imaged	orthogonal	three	axes	degree	speed	degrees	orientation
Topic 46	volume	phantom	volumes	homogeneous	distance	investigation	rendering	filled	partial	spatially
Topic 47	loop	circuit	capacitor	resonant	tuning	loops	parallel	impedance	capacitors	decoupling
Topic 48	sensitivity	receiving	reduced	coils	parallel	reconstruction	received	fov	profiles	receiver
Topic 49	device	medical	devices	invasive	microcoil	procedures	microcoils	adapted	components	response
Topic 50	treatment	breast	therapy	lesion	radiation	dose	microscope	localization	lesions	prostate
Topic 51	ring	rings	shielding	spaced	annular	radially	turns	structure	axially	tubular
Topic 52	temperature	cooling	heat	thermal	cell	heating	coolant	cooled	cryogenic	thermally
Topic 53	patient	base	joint	table	bed	head	positioning	movement	neck	mounted
Topic 54	element	sensor	zone	anterior	posterior	vicinity	interrogation	contact	mass	adapted
Topic 55	particles	carrier	coating	particle	solid	liquid	aqueous	ferromagnetic	crystal	material
Topic 56	optical	transducer	ultrasonic	fiber	patient	sound	monitor	sensor	electrical	fibers
Topic 57	cycle	segment	segments	cardiac	ecg	respiratory	100	cycles	110	102
Topic 58	computer	images	ultrasound	diagnosis	anatomical	medical	parameter	analysis	program	color
Topic 59	location	marker	orientation	internal	locations	markers	physical	virtual	structure	fiducial
Topic 60	current	eddy	currents	compensation	electric	induced	compensating	paths	compensate	flow
Topic 61	agents	metal	contrast	paramagnetic	compounds	complexes	diagnostic	ions	acid	mri
Topic 62	samples	test	analysis	measurements	curve	mass	testing	analyte	capillary	log
Topic 63	antenna	transmission	radio-frequency	reception	transmitting	receiving	antennas	local	auxiliary	examination
Topic 64	compound	compounds	molecule	labeled	mixture	molecules	identifying	protein	ligand	detecting
Topic 65	k-space	sampling	partial	dynamic	trajectory	ssfp	reconstruction	spiral	central	sampled
Topic 66	point	points	intensity	map	grid	boundary	contour	distance	contours	seed
Topic 67	nmr	nuclear	experiment	measurement	spectroscopy	spin-echo	spectrometer	detection	techniques	nutation
Topic 68	circuit	receiver	voltage	transmitter	channel	switch	switching	circuitry	channels	power
Topic 69	medical	user	interface	diagnostic	remote	service	software	network	database	stored
Topic 70	excitation	pulses	profile	nuclear	radio-frequency	tomography	sequence	excited	readout	dataset
Topic 71	tube	bore	vacuum	magnet	superconducting	vessel	assembly	thermal	shield	superconductive
Topic 72	spatial	series	spatially	temporal	coefficients	distribution	linear	modulation	limit	exposure
Topic 73	low	large	mri	performance	reduced	loss	bandwidth	artifact	makes	easy
Topic 74	axis	longitudinal	central	parallel	transverse	perpendicular	oriented	plane	distance	angle
Topic 75	tissue	delivery	tumor	tissues	soft	energy	surrounding	ablation	muscle	patient
Topic 76	needle	specimen	biopsy	3-d	cannula	cutting	cap	spring	device	border

Topic 77	model	matrix	density	continuous	algorithm	solution	quantity	coefficient	rows	linear
Topic 78	power	vessel	liquid	helium	storage	tank	refrigerator	source	magnet	dewar
Topic 79	motion	reconstruction	reconstructed	images	artifacts	frames	velocity	mri	series	moving
Topic 80	contrast	agent	medium	bolus	administering	administered	agents	concentration	effective	after
Topic 81	three-dimensional	memory	view	representation	processor	displayed	stored	screen	operator	images
Topic 82	formula	atoms	alkyl	hydrogen	str1	compounds	carbon	c.sub.1	atom	compound
Topic 83	antibiotic	streptomyces	compound	fermentation	antibiotics	animals	medium	compounds	antibacterial	strain
Topic 84	energy	electromagnetic	operation	wave	acoustic	waves	r.f	microwave	emitted	noise
Topic 85	elements	center	element	sections	planar	peripheral	shape	array	curved	pair
Topic 86	mri	scanner	room	adapted	shielded	magnet	configured	scan	positioning	receive
Topic 87	inner	outer	segments	wall	cylinder	cylindrical	exterior	mounted	interior	surface
Topic 88	line	lines	pattern	transmission	distance	feeding	connecting	length	patterns	fuel
Topic 89	complex	scanning	phase	scanned	encoded	conjugate	utilizing	incomplete	unique	corrected
Topic 90	resonator	head	conductive	electrically	dielectric	conducting	substrate	resonators	loops	extension
Topic 91	surface	structure	side	upper	geometry	flat	board	sides	boundary	front
Topic 92	conductor	conductors	cable	electrically	coaxial	electrical	connector	connecting	length	surface
Topic 93	housing	plates	rotor	face	surfaces	cylindrical	pole	closed	circular	shape
Topic 94	electrical	generator	electrode	lead	electrodes	sensing	device	implantable	medical	contact
Topic 95	static	detection	detecting	detected	nuclear	generation	detect	detects	inspection	applies
Topic 96	unit	controller	units	communication	controls	receiving	transmitting	generates	controlling	receives
Topic 97	shield	conductive	electrically	dielectric	sheet	currents	material	strip	strips	eddy
Topic 98	material	layer	layers	metal	thin	materials	alloy	strength	non-magnetic	wire
Topic 99	support	frame	supporting	table	plate	fixed	vertical	horizontal	base	opening
Topic 100	component	components	response	interference	separated	noise	pulses	received	generation	buffer

Table A2- Top 10 terms of the identified 100 topics in abstracts of RFID patents

Topic 1	conductive	chip	electrically	dielectric	pattern	substrate	such	ink	conductor	material
Topic 2	carrier	modulation	modulated	phase	backscatter	modulator	amplitude	demodulator	baseband	demodulation
Topic 3	sensor	temperature	monitoring	environmental	sensors	measurement	sensing	food	induction	ambient
Topic 4	location	locations	tracking	such	locating	geographic	map	environment	database	real-time
Topic 5	tracking	track	technology	tagged	readers	inventory	such	portal	configured	tracked
Topic 6	product	database	rack	unique	epc	scanner	placement	such	consumer	manufacturer
Topic 7	memory	stored	storing	cell	operation	non-volatile	write	coded	writing	radio-frequency
Topic 8	layer	film	metal	layers	protective	material	thin	paper	plastic	aluminum
Topic 9	transaction	account	consumer	payment	transactions	identifier	financial	point-of-sale	sale	merchant
Topic 10	substrate	integrated	flexible	array	contact	glass	electrically	substrates	capsule	forming
Topic 11	store	personal	document	scanner	documents	apparatus	such	incorporated	software	integrated
Topic 12	resonant	coil	clock	capacitor	resonance	inductive	parallel	spiral	amplifier	frequencies
Topic 13	network	local	node	nodes	auto-id	address	architecture	wlan	distributed	such
Topic 14	code	bar	codes	unique	stored	such	operation	manufactured	electronic	password
Topic 15	communication	terminal	address	aircraft	terminals	session	indicating	guiding	receiving	self-service
Topic 16	apparatus	panel	animal	contactless	unauthorized	disable	point	disabled	animals	controlling
Topic 17	transponder	transponders	programming	toy	communicate	programmable	adapted	wireless	shield	programmer
Topic 18	voltage	current	transistor	capacitor	semiconductor	reference	gate	output	source	switching
Topic 19	power	battery	source	transmitted	electric	low	operation	powered	consumption	reduced
Topic 20	detection	detector	generator	proximity	detected	excitation	generates	detects	detect	response
Topic 21	side	board	attachment	printed	strip	flexible	chip	inlay	mounted	such
Topic 22	surface	adhesive	pattern	layer	wear	face	thickness	release	secured	microchip
Topic 23	external	passive	logic	internal	port	circuits	activation	chip	such	ports
Topic 24	adapted	asset	communicate	operable	monitor	monitoring	communication	management	store	wirelessly
Topic 25	package	light	indicator	visual	indication	exposure	shelf	shelves	visible	slide
Topic 26	service	provider	network	resource	services	internet	registered	audio	resources	record
Topic 27	antennas	energy	source	electrical	polarized	dipole	such	reflected	tuned	diode
Topic 28	computer	program	software	automatically	hardware	physical	reporting	query	such	enabled
Topic 29	central	parking	current	maintenance	vehicles	baggage	management	database	lot	aircraft
Topic 30	article	articles	eas	surveillance	temporary	electronic	household	inventory	stored	book
Topic 31	electronic	such	integrated	included	circuits	allow	incorporates	fabrication	volatile	approaches
Topic 32	components	electrical	manufacturing	contacts	organic	such	manufacture	component	contact	cost
Topic 33	label	labels	closure	liner	printed	removable	removed	sheet	after	sealing
Topic 34	host	printer	computer	interface	such	register	color	cash	protocol	modular
Topic 35	member	structure	chamber	extending	such	driven	head	air	surface	side
Topic 36	media	mechanism	door	locking	compartment	lock	work	signature	housing	removed
Topic 37	inventory	gps	cart	positioning	satellite	shopping	global	communications	division	location

Topic 38	interrogation	interrogating	readers	response	units	noise	signals	interrogated	detecting	verifier
Topic 39	vehicle	driver	vehicles	motor	mounted	seat	installed	speed	doors	road
Topic 40	authentication	key	random	secure	encrypted	encryption	keys	request	authenticating	valid
Topic 41	physical	environment	platform	facility	distribution	computing	locations	virtual	infrastructure	building
Topic 42	module	interface	modules	subsystem	indicating	interfaces	microcontroller	microprocessor	usb	operation
Topic 43	security	sensors	such	communications	proximity	external	contain	user	communicate	behavior
Topic 44	housing	plate	metal	slot	assembly	wall	holder	front	cavity	interior
Topic 45	customer	weight	party	purchase	price	telephone	customers	checkout	discount	purchaser
Topic 46	storage	storing	stored	transfer	inspection	serial	such	retrieval	sound	limited
Topic 47	server	management	collection	updated	network	processed	stores	transmits	moveable	middleware
Topic 48	mail	delivery	shipping	goods	shipment	sorting	recipient	destination	such	bin
Topic 49	printing	identifying	cable	equipment	search	such	installed	identify	cables	stolen
Topic 50	optical	image	imaging	images	capture	scanning	linking	such	handling	x-ray
Topic 51	battery	assembly	web	chips	charging	source	solid	reel	heat	pack
Topic 52	mobile	dynamic	receives	phone	telephone	representation	call	center	implementation	short-range
Topic 53	card	token	purchase	business	smart	cards	credit	purchasing	merchant	consumer
Topic 54	medical	automatically	operator	prescription	instructions	such	records	pharmacy	worker	processor
Topic 55	tape	core	aperture	fiber	cartridge	surface	length	reinforcing	wall	shaped
Topic 56	interrogator	exciter	interrogators	operational	values	reflective	sequence	commands	frequencies	intended
Topic 57	container	containers	tray	cap	contents	such	continuously	mounted	reusable	contained
Topic 58	format	barcode	xml	stream	action	rules	receives	manager	formats	template
Topic 59	portable	electrical	connector	enclosure	receptacle	electrically	socket	activated	configured	plug
Topic 60	patient	profile	authorization	stored	limited	medication	sales	shopper	merchandise	activate
Topic 61	identifier	unique	identifiers	cartridge	continuous	read	photo	stored	facility	motion
Topic 62	person	destination	biometric	call	elevator	room	personnel	mechanism	floor	movement
Topic 63	impedance	line	matching	coupler	transmission	lines	conductor	strip	length	load
Topic 64	magnetic	switch	coil	devices	coils	charge	transport	materials	optical	electronic
Topic 65	received	smart	performance	strength	cellular	pallet	such	configured	values	jamming
Topic 66	remote	digital	analog	link	command	internet	non-contact	operation	decoding	appliances
Topic 67	machine	check	readable	payment	cards	such	vending	adapted	atm	operative
Topic 68	elements	element	switching	electronics	such	assembly	tank	hole	inner	pair
Topic 69	electromagnetic	wave	acoustic	waves	electric	pump	cassette	radiation	transducer	intensity
Topic 70	gaming	game	play	table	calibration	player	clip	participants	ride	players
Topic 71	print	band	printing	mechanism	printer	uhf	printed	head	release	cutting
Topic 72	sheet	support	dispenser	dispensing	wafer	flow	beverage	semiconductor	thermal	inlet
Topic 73	message	remote	program	voice	interaction	messages	client	modem	ball	response
Topic 74	test	fluid	testing	tester	segment	apparatus	biological	segments	read/write	measurement
Topic 75	reader/writer	embedded	recorded	disc	optical	paper	writing	encoded	annular	written
Topic 76	image	tire	forming	tamper	pressure	wheel	bag	resistant	distance	bags

Topic 77	tool	channel	transceivers	short-range	accessory	communication	locator	channels	tools	broadcast
Topic 78	material	materials	packaging	surrounding	hazardous	probe	handling	absorbing	contained	bulk
Topic 79	base	station	food	transmitted	entity	promotional	proximity	offers	safety	communicated
Topic 80	circuitry	sensing	alarm	electrostatic	sensed	movement	opening	activating	sense	lid
Topic 81	instrument	surgical	pin	identifying	fob	fuel	instruments	transit	counter	count
Topic 82	seal	messages	badge	cargo	orientation	thread	lifting	monitoring	sealed	such
Topic 83	zone	frame	scanning	indicia	zones	license	scanned	failure	markers	door
Topic 84	video	monitoring	such	camera	tracking	cameras	surveillance	people	locating	facilitate
Topic 85	loop	mhz	frequencies	modulated	outbound	integrated	inbound	negative	figure	operable
Topic 86	medium	recording	edge	cover	frame	transmitting	apparatus	window	accommodating	outer
Topic 87	wireless	communication	communications	devices	such	network	communicate	networks	point	protocol
Topic 88	user	users	input	exercise	personal	name	dynamic	list	file	kiosk
Topic 89	signals	received	receive	receiving	receives	conveyor	transmit	responsive	array	belt
Topic 90	response	transmitted	received	command	request	transmits	identifying	receives	receipt	bits
Topic 91	element	plane	patch	ground	feed	radiating	radiation	point	feeding	angle
Topic 92	receiver	transmitter	input	output	receivers	converter	feedback	receives	receive	decoder
Topic 93	component	transceiver	controller	receive	communication	communicates	such	components	facilitates	industrial
Topic 94	devices	fixed	such	protocols	pointer	implementations	able	operation	interference	standard
Topic 95	configured	processor	path	receive	point	transmit	corrugated	travel	send	operation
Topic 96	read	reading	configuration	machine-readable	write	such	feedback	reads	written	readers
Topic 97	receiving	transmitting	parameter	status	detecting	received	storing	sending	controlling	identifying
Topic 98	unit	trigger	units	receives	receiving	load	communicates	identifies	sending	determines
Topic 99	transmission	transmit	reception	reply	receive	reference	operate	techniques	bandwidth	communications
Topic 100	detected	detecting	cycle	verifying	verify	upper	closed	opening	read	detects

Table A3- Top 10 terms of the identified 100 topics in abstracts of Stem Cell patents

Topic 1	tumor	cancer	telomerase	tumors	solid	mammary	compositions	treatment	inhibit	ability
Topic 2	tissues	living	periodontal	bioactive	regeneration	injured	implanted	repair	configuration	anchor
Topic 3	formula	compounds	atom	compound	alkyl	hydrogen	represented	r.sup.1	derivative	salt
Topic 4	liver	hepatic	hepatocytes	progenitors	primitive	isolated	osteoblasts	tissues	proximal	hepatocyte
Topic 5	surface	lines	media	grown	increasing	toxicity	membrane	expressing	adhesion	density
Topic 6	diseases	disorders	treating	diabetes	treat	disease	genetic	metabolic	genes	procedure
Topic 7	molecules	conjugates	compositions	mammalian	kits	formation	remodeling	moiety	fragments	modifying
Topic 8	dna	sequence	promoter	nucleotide	encoding	linked	artificial	chromosomes	regulatory	rna
Topic 9	includes	barrier	support	porous	plasma	liquid	apparatus	components	product	fluid
Topic 10	proteins	chimeric	domain	fusion	signal	receptors	expressing	cytoplasmic	fragments	binding
Topic 11	system	nervous	central	treatment	cns	reducing	screening	myelin	specifically	diseases
Topic 12	embryos	primordial	porcine	cloning	nuclear	animals	embryonic	germ	culture	transfer
Topic 13	protein	kinase	tyrosine	compositions	specifically	receptor	hscs	ligands	sequences	involved
Topic 14	genetic	material	incorporated	male	biocompatible	germ	transfected	exogenous	vitro	three-dimensional
Topic 15	embryonic	undifferentiated	pluripotent	differentiate	differentiated	ability	proliferate	hemopoietic	feeder	mature
Topic 16	oligonucleotide	three	molecule	bonds	compound	synthetic	dimensional	elements	residues	forming
Topic 17	dna	recombination	homologous	chromosomal	targeting	loci	eukaryotic	genetic	embryonic	genomic
Topic 18	binding	donor	bacteria	bacterial	bind	compounds	proliferative	large	alpha	screening
Topic 19	blood	peripheral	cord	umbilical	after	red	marrow	bag	monocytes	white
Topic 20	factor	stimulating	colony	epidermal	granulocyte	interleukin	hepatocyte	insulin-like	fibroblast	cytokine
Topic 21	skin	equivalent	basal	layer	media	epidermis	epithelial	vitro	organotypic	feeder
Topic 22	immune	response	dendritic	system	organs	rejection	development	transplanted	suppressing	suppress
Topic 23	transgenic	animals	non-human	animal	expression	mammals	transgene	nonhuman	altered	receptor
Topic 24	culture	adherent	grow	carriers	medium	cultivating	harvested	cultivated	extracellular	feeder
Topic 25	biological	stimulation	electrical	polymer	cellular	polymers	substances	devices	electromagnetic	conductive
Topic 26	genetically	engineered	express	genetic	response	encapsulated	improving	delivery	encoding	modification
Topic 27	compounds	chemokine	promoting	receptor	analogues	progenitor	compositions	lipid	hiv	maturation
Topic 28	host	recipient	graft	transplant	donor	administering	disease	versus	malignant	transplantation
Topic 29	vivo	vitro	compositions	culture	enhance	proliferating	expansion	hematopoiesis	stimulating	stimulation
Topic 30	embryonic	drug	embryoid	differentiation	lineage	markers	screening	culture	bodies	differentiated
Topic 31	mammal	administering	treating	damaged	effective	inducing	therapeutically	failure	isolated	implanting
Topic 32	heart	devices	myocardium	cardiac	myocardial	manipulation	living	sheets	culture	promote
Topic 33	expression	genes	therapy	foreign	proteins	heterologous	transformed	encoded	constructs	transfected
Topic 34	fetal	role	line	positive	embryonic	germ	marker	culture	play	induced
Topic 35	isolated	retinal	vivo	surface	propagation	progenitor	assays	markers	developing	differentiation
Topic 36	oxygen	mammalian	anchors	oocyte	allow	substrate	vitro	construct	tension	cultured
Topic 37	mouse	mice	mutant	transgenic	model	disruption	homozygous	null	knockout	embryonic

Topic 38	endothelial	vascular	hematopoiesis	primate	cryopreserved	isolated	tissues	factor	adult	transplanted
Topic 39	scaffold	precursor	biocompatible	ligament	matrix	core	jacket	external	mechanical	forming
Topic 40	neurons	neuronal	neural	brain	treating	neurological	dopaminergic	disorders	neuron	disease
Topic 41	progenitor	cd34.sup	primitive	marrow	isolated	myeloid	endothelial	cd38.sup	purified	cytokines
Topic 42	differentiation	proliferation	development	inducing	induce	vitro	lineage	survival	lineages	cellular
Topic 43	receptor	ligand	peptide	agonist	binds	antagonist	lytic	soluble	extracellular	domain
Topic 44	population	differentiation	contacting	expansion	effective	inhibitor	increasing	inducing	thrombopoietin	proliferation
Topic 45	expression	lung	death	involving	deficiency	disorders	activities	analysis	apoptosis	affecting
Topic 46	antibody	marker	fragment	fragments	antigen	mesenchymal	dsm	identifying	progenitor	hybridoma
Topic 47	catheter	delivery	balloon	injection	rotatable	stent	assembly	lumen	vein	implantation
Topic 48	sequence	vector	genome	cassette	heterologous	site-specific	library	inserted	sequences	includes
Topic 49	device	vessel	lumen	includes	expandable	medical	apparatus	assembly	member	delivery
Topic 50	leukemia	chronic	acute	lymphoma	diagnosing	develop	found	lymphoblastic	myelogenous	clinical
Topic 51	acid	nucleic	amino	acids	sequence	encoding	polypeptide	seq	isolated	residue
Topic 52	agent	cytotoxic	administering	therapy	effective	cd4	chemotherapy	treatment	therapeutic	vivo
Topic 53	culture	vitro	differentiated	pancreatic	expanded	adult	differentiation	cultures	mature	diabetes
Topic 54	layer	component	scaffold	matrix	calcium	phosphate	dermal	extracellular	carrier	biodegradable
Topic 55	wound	healing	system	components	transfer	transferrin	direct	protein	treatment	skin
Topic 56	hair	skin	removal	removing	keratinocytes	modulating	detecting	follicle	agent	light
Topic 57	peptides	polynucleotide	whole	viruses	peptide	concerns	stimulate	recombinant	linear	identical
Topic 58	antibodies	antigen	prostate	cancer	monoclonal	detecting	antibody	humanized	bind	suspensions
Topic 59	blood	vessels	antigens	mononuclear	flow	ischemic	diseased	coronary	increasing	become
Topic 60	cartilage	repair	chondrocytes	connective	injury	oxide	defects	nitric	collagen	shear
Topic 61	membrane	storage	amniotic	includes	container	filter	covering	making	mechanism	storing
Topic 62	implant	defect	particles	biocompatible	pores	porous	material	shaped	demineralized	magnetic
Topic 63	resistance	lymphocytes	t-cell	receptors	tumor	multidrug	humans	t-cells	resistant	antigen
Topic 64	structure	graft	stromal	preparing	prosthetic	implant	surface	three-dimensional	substrates	aggregated
Topic 65	injury	cord	spinal	model	brain	dose	activated	rodent	effective	therapeutic
Topic 66	animal	alpha	endogenous	plant	animals	prions	physiological	carbohydrate	sugar	assay
Topic 67	neural	multipotent	progeny	neurons	precursor	oligodendrocytes	cns	astrocytes	glial	nscs
Topic 68	solution	solutions	biological	freezing	exposed	stabilizing	network	temperature	nutrient	concentration
Topic 69	cardiac	organ	system	heart	muscle	implantable	myocardial	skeletal	therapy	repair
Topic 70	vertebrate	expression	genome	lambda	mouse	viral	hiv-1	transformed	downstream	mutants
Topic 71	mesenchymal	serum	culturing	electric	isolated	effective	induction	forming	includes	essential
Topic 72	bioreactor	hollow	fiber	placenta	residual	placental	perfusion	remove	liquid	gel
Topic 73	marrow	transplantation	blood	stromal	therapy	autologous	peripheral	sites	hemopoietic	flow
Topic 74	epithelial	phenotype	after	colony	colonies	expression	altered	genetically	medium	enrichment
Topic 75	scf	analogs	erythroid	medium	recombinant	disclosure	factor	epo	g-csf	gm-csf
Topic 76	protein	proteins	recombinant	encoding	family	isolated	expressing	purified	host	concerns

Topic 77	mhc	swine	species	genes	antigens	antigen	genetic	mature	targeting	tolerance
Topic 78	mixture	separation	particles	complex	suspension	heterogeneous	lymphocyte	separated	purging	contacted
Topic 79	treating	preventing	compounds	modulating	expression	compositions	prevent	identifying	inventors	found
Topic 80	vectors	polypeptides	sequences	polynucleotides	expression	encoding	compositions	dna	plant	regulatory
Topic 81	vector	viral	retroviral	virus	vectors	particles	packaging	hiv	recombinant	transfer
Topic 82	binding	polypeptide	assays	polypeptides	recombinant	affinity	encoding	retrovirus	domain	altered
Topic 83	compositions	pharmaceutical	kits	administration	pharmaceutically	carrier	delivery	administering	preparations	preparing
Topic 84	treatment	administration	therapy	prevention	patient	therapies	marrow	effective	g-csf	cancer
Topic 85	agents	screening	therapeutic	biological	potential	drugs	agent	preparing	identify	therapies
Topic 86	mammalian	construct	constructs	genes	identification	exogenous	concerns	dna	isolation	genome
Topic 87	member	wall	surface	outer	inner	femoral	shaped	side	bore	mounted
Topic 88	beta	differentiation	catenin	signaling	alpha	tgf	sub.2	transforming	hedgehog	chain
Topic 89	material	substrate	surface	location	artificial	matrix	support	biologically	energy	receiving
Topic 90	patient	administered	patients	after	adipose	treating	suffering	autologous	procedure	vivo
Topic 91	line	embryonic	lines	stage	cardiomyocytes	differentiate	stable	embryo	establishing	blastocyst
Topic 92	fluid	flow	first	chamber	second	particles	system	stent	liquid	separating
Topic 93	culture	medium	culturing	cultured	stromal	conditioned	avian	feeder	free	mammalian
Topic 94	population	populations	enriched	isolating	progenitor	identifying	mixed	fluorescent	marker	myeloid
Topic 95	factor	fibroblast	muscle	test	ependymal	lif	compound	inhibitory	proliferation	leukemia
Topic 96	therapeutic	treatment	disorders	diseases	diagnostic	compositions	patients	prevention	protocols	prophylactic
Topic 97	weight	molecular	regeneration	compositions	hydrogel	acid	medical	promote	activation	hyaluronic
Topic 98	disease	treatment	disorder	administering	transplantation	alzheimer	includes	neurodegenerative	treating	cerebral
Topic 99	reporter	transcription	element	factor	identifying	promoter	enhancer	marker	expresses	isolating
Topic 100	matrix	extracellular	matrices	fibers	materials	natural	surfaces	repair	biological	scaffolds

Table A4- Top 10 terms of the identified 100 topics in abstracts of Nanotube patents

Topic 1	nanoparticles	complex	binding	tissue	diffraction	biological	grating	form	medical	activity
Topic 2	electrode	cnt	cnts	formed	carbon	disposed	connected	electrically	device	current
Topic 3	component	core	inorganic	cover	anisotropic	comprises	golf	flexible	material	ball
Topic 4	single	precursor	wall	single-walled	carbon	nanotubes	support	selected	ni	si
Topic 5	groups	functional	chemical	bonded	functionalized	chemically	group	modified	moiety	species
Topic 6	light	emitting	actuator	receiving	pixel	photosensitive	wavelength	fluorescent	visible	portions
Topic 7	located	conductor	cable	connecting	connected	housing	central	connection	connector	vehicles
Topic 8	gas	hydrogen	pressure	storage	vessel	temperature	storing	stream	adsorption	adsorbent
Topic 9	electrolyte	solid	electrochemical	cell	electrodes	diffusion	ions	battery	cells	regeneration
Topic 10	sensor	sensing	frequency	sensors	analyte	change	detecting	signal	antenna	detection
Topic 11	vapor	deposition	chamber	chemical	gas	grown	growth	carbon	catalytic	source
Topic 12	conducting	monolayer	comprises	tunneling	method	preparing	presence	self-assembled	substances	brought
Topic 13	control	switching	element	output	signal	input	node	channel	switch	signals
Topic 14	heat	thermal	interface	transfer	thermally	sink	cooling	chip	die	material
Topic 15	source	drain	gate	terminal	nanotube	coupled	field	terminals	electrically	electrical
Topic 16	devices	methods	nanostructures	nanowires	nanostructure	directed	electronic	making	systems	techniques
Topic 17	magnetic	base	field	recording	head	resonance	medium	measuring	current	operation
Topic 18	method	manufacturing	fabricating	fabrication	cost	nano-tube	utilizing	process	simple	dispersing
Topic 19	polymer	polymers	compositions	relates	monomer	prepared	polymerization	aromatic	composition	blend
Topic 20	discharge	carbon	arc	gas	plasma	apparatus	inert	generated	soot	graphite
Topic 21	layer	formed	dielectric	hole	patterned	intermediate	layers	overlying	sidewall	sacrificial
Topic 22	electronemitting	glass	grid	electron	electrode	phosphor	substrate	member	source	Space
Topic 23	resin	wt	filler	thermoplastic	curable	resist	cured	molded	impact	comprises
Topic 24	semiconductor	device	nanowire	doped	transistors	dopant	thin-film	method	n-type	vertical
Topic 25	cathode	anode	cold	vacuum	field	x-ray	electrons	device	electric	current
Topic 26	gate	formed	insulating	cathode	layer	electrode	emission	substrate	insulation	hole
Topic 27	field	emission	plate	electric	display	emitters	emitting	device	formed	applied
Topic 28	region	transistor	channel	gate	device	structure	drain	nanotube	trench	floating
Topic 29	probe	scanning	portion	cantilever	microscope	nanotube	force	base	holder	needle
Topic 30	circuit	nanoscale	integrated	output	input	programmable	wire	coupled	logic	data
Topic 31	fluid	fuel	water	cell	oxygen	hydrogen	supercritical	device	membrane	communication
Topic 32	display	panel	assembly	transparent	crystal	flat	illumination	plasma	lens	disposed
Topic 33	conductive	electrically	electrical	ribbon	non-conductive	circuit	electromechanical	traces	circuits	disposed
Topic 34	material	capacitor	characteristics	activated	battery	lithium	nanoelement	carbon-based	alkali	bodies
Topic 35	active	improved	efficiency	conversion	energy	reduced	achieved	performance	addition	improve
Topic 36	electrical	contact	resistance	conductors	wiring	contacts	wires	connection	exposure	device
Topic 37	emitter	paste	protective	electron	field	layer	negative	vacuum	mask	screen

Topic 38	beam	radiation	ion	image	detector	x-ray	ions	source	infrared	incident
Topic 39	film	films	substrate	formed	thickness	adhesion	shield	typically	adherent	nanostructured
Topic 40	liquid	thermal	conductivity	medium	expansion	dispersion	fluid	pipe	coefficient	dispersed
Topic 41	portion	opening	formed	exposed	surface	photoresist	portions	nozzle	spray	etching
Topic 42	nanotubes	carbon	semiconducting	metallic	formed	ends	nanotube	selectively	utilized	create
Topic 43	electron	emission	electrons	emitting	emitted	device	beams	sources	vacuum	focusing
Topic 44	catalyst	carbon	growth	gas	metal	grow	grown	reactor	catalytic	growing
Topic 45	molecules	molecule	methods	dna	probes	mass	analysis	detection	nucleic	acid
Topic 46	molecular	network	physical	neural	based	species	gap	lattice	connections	set
Topic 47	composite	matrix	fiber	composites	material	ceramic	reinforced	friction	dispersed	reinforcement
Topic 48	organic	electronic	compositions	devices	photoactive	making	polymeric	donor	photovoltaic	aqueous
Topic 49	array	aligned	object	provide	objects	vertically	ordered	nano-sized	nanotube-based	arrays
Topic 50	reaction	proton	conductor	mixture	product	zone	reacting	derivative	intermediate	reactant
Topic 51	apparatus	heating	time	plasma	electromagnetic	method	gaseous	methods	heated	microwave
Topic 52	laser	Target	beam	selected	method	energy	irradiation	reagent	evaporation	evaporated
Topic 53	electrodes	formed	electrode	predetermined	emitters	therebetween	gap	pair	phosphor	opposing
Topic 54	coating	coated	binder	material	formed	coatings	shielding	printing	glass	comprised
Topic 55	memory	cell	cells	nano	bit	word	device	read	stored	plurality
Topic 56	energy	charge	quantum	transport	electric	charging	devices	fields	systems	thermoelectric
Topic 57	optical	storage	data	medium	media	absorption	energy	device	comprises	wave
Topic 58	arranged	upper	plurality	openings	holes	display	lines	spacer	cathodes	formed
Topic 59	fullerene	group	substituted	formula	derivatives	derivative	c60	alkyl	independently	groups
Topic 60	temperature	range	room	enhanced	rate	transition	temperatures	wide	annealing	suitable
Topic 61	fullerenes	compounds	fullerene	starting	tubular	treated	soluble	solvents	relates	cleaning
Topic 62	surface	surfaces	interface	extend	silver	grooves	colloid	embedded	portions	uniformly
Topic 63	process	plasma	formation	etching	prior	production	purification	producing	preparation	processes
Topic 64	direction	oriented	parallel	axis	perpendicular	orientation	graphite	plane	longitudinal	sheet
Topic 65	metal	oxide	catalytic	transition	alloy	titanium	iron	carbide	oxides	nickel
Topic 66	body	formed	cavity	treatment	form	superconducting	aluminum	lead	optionally	cell
Topic 67	layers	form	formed	individual	multilayer	layer	layered	multiple	stack	colorant
Topic 68	fibers	filter	strength	product	processes	provide	conventional	application	concentration	environment
Topic 69	material	property	materials	fibrous	melting	capable	embedded	preparing	superior	bent
Topic 70	method	producing	step	comprises	relates	steps	produced	carbon	oxidizing	production
Topic 71	agent	agents	substance	composition	modifier	combination	dispersion	contrast	pharmaceutical	amino
Topic 72	form	diamond	graphite	mixture	powder	amorphous	crystalline	carbon-containing	carbon	atmosphere
Topic 73	element	elements	member	electric	members	connected	structural	formed	electrically	applied
Topic 74	structure	form	plural	mesh	method	fed	point	forming	applying	nano-sized
Topic 75	article	defined	aspect	fabric	polishing	articles	suspended	segments	embodiments	methods
Topic 76	phase	porous	membrane	continuous	synthesis	reactions	methods	oxidation	chemical	sites

Topic 77	materials	properties	methods	mechanical	electrical	suitable	physical	combinations	applications	processing
Topic 78	nanotube	carbon	quartz	pair	causing	contacted	flowing	bundles	etched	microelectromechanical
Topic 79	area	nanofibers	controlled	predetermined	nanofiber	surface	small	shape	large	desired
Topic 80	nm	diameter	metallic	hollow	size	outer	average	diameters	pore	bearing
Topic 81	carbon	furnace	excellent	nano-size	produced	network	subjecting	ranges	modified	flat
Topic 82	weight	mixture	particulate	parts	particulates	mass	relates	ratio	percent	nanocomposite
Topic 83	swnt	applications	arrays	nanotubes	swnts	variety	embodiment	ends	growth	nanotube
Topic 84	unit	power	type	supply	secondary	primary	units	transmission	time	generated
Topic 85	ratio	volume	pulse	filament	long	distance	bundle	surface	pulses	major
Topic 86	substrate	pattern	formed	deposited	patterned	predetermined	substrates	surface	grown	thereon
Topic 87	device	voltage	applied	current	potential	driving	applying	response	level	threshold
Topic 88	embodiment	support	regions	comprise	preferred	attached	comprises	device	band	spin
Topic 89	fullerenes	c60	solid	soot	hydrocarbon	combustion	nanomaterials	flame	zone	carbon
Topic 90	single-wall	carbon	nanotubes	silicon	nitride	mold	boron	relates	form	carbide
Topic 91	forming	method	providing	substrate	steps	depositing	removing	form	making	patterning
Topic 92	components	polymeric	block	constant	copolymer	diamondoid	higher	ceramic	relates	diamondoids
Topic 93	particles	particle	template	fine	form	carrier	size	pores	alumina	dispersed
Topic 94	preferably	length	greater	free	equal	shaped	mum	density	small	set
Topic 95	sample	mode	tips	separation	force	specific	comprises	capillary	measurement	features
Topic 96	solution	solvent	acid	suspension	form	dissolved	dispersion	solutions	method	aqueous
Topic 97	plurality	dielectric	disposed	adjacent	comprises	edge	position	selectively	media	controllably
Topic 98	structures	comprises	nanoparticle	formation	embodiments	hybrid	method	embodiment	dispersing	provide
Topic 99	compound	group	selected	metal	formula	contacting	method	transition	linear	metals
Topic 100	atoms	carbonaceous	tube	carbon	material	formed	cluster	clusters	cylindrical	composed

Table A5: The moderating role of application breadth

Estimation Model:	Logistic (Odds Ratios Reported)	
DV:	Novel Breakthrough	Economic Breakthrough
	(1)	(2)
Novel breakthroughs		1.366*** (0.003)
First inventor's knowledge breadth ×		
Narrowly applied	1.271 (P=0.108)	1.045* (P=0.091)
Broadly applied	1.001 (P=0.990)	1.238*** (P=0.000)
First inventor's knowledge depth ×		
Narrowly applied	0.843* (P=0.053)	1.061 (P=0.128)
Broadly applied	0.860 (P=0.712)	0.945*** (P=0.000)
Second inventor's knowledge breadth ×		
Narrowly applied	0.854* (P=0.080)	1.012 (P=0.780)
Broadly applied	1.549*** (P=0.001)	0.955 (P=0.515)
Second inventor's knowledge depth ×		
Narrowly applied	1.032 (P=0.554)	1.028*** (P=0.000)
Broadly applied	0.552** (P=0.046)	0.922 (P=0.318)
Overlap in the knowledge scope of the first and second inventors ×		
Narrowly applied	0.699 (P=0.680)	1.020 (P=0.827)
Broadly applied	0.979 (P=0.864)	0.950 (P=0.455)
Third inventor's knowledge breadth ×		
Narrowly applied	0.942* (P=0.060)	0.960 (P=0.087)
Broadly applied	1.259 (P=0.133)	1.034 (P=0.780)
Third inventor's knowledge depth ×		
Narrowly applied	1.438 (P=0.164)	0.992 (P=0.594)
Broadly applied	0.772 (P=0.234)	1.185** (P=0.018)

Fourth inventor's knowledge breadth ×		
Narrowly applied	0.675*** (P=0.000)	1.061 (P=0.521)
Broadly applied	0.636** (P=0.025)	1.023 (P=0.795)
Fourth inventor's knowledge depth ×		
Narrowly applied	1.490*** (P=0.000)	0.958 (P=0.478)
Broadly applied	1.803*** (P=0.000)	0.901 (P=0.023)
Fifth inventor's knowledge breadth ×		
Narrowly applied	0.887*** (P=0.000)	0.961 (P=0.109)
Broadly applied	1.020 (P=0.944)	0.927 (P=0.361)
Fifth inventor's knowledge depth ×		
Narrowly applied	0.823*** (P=0.002)	1.034 (P=0.580)
Broadly applied	1.013 (P=0.950)	1.039* (P=0.062)
Full set of controls	Yes	Yes
Technology-time fixed effects	Yes	Yes
Number of observations	4,110	16,905
R-squared	0.279	0.185

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table A6: The moderating role of technological modularity

Estimation Model:	Logistic (Odds Ratios Reported)	
DV:	Novel Breakthrough	Economic Breakthrough
	(1)	(2)
Novel breakthroughs		1.324**
		(0.015)
First inventor's knowledge breadth ×		
Non-Modular	1.328**	1.229***
	(P=0.018)	(P=0.000)
Modular	0.925***	1.052*
	(P=0.000)	(P=0.098)
First inventor's knowledge depth ×		
Non-Modular	0.724***	0.875***
	(P=0.000)	(P=0.000)
Modular	0.995	1.063**
	(P=0.946)	(P=0.035)
Second inventor's knowledge breadth ×		
Non-Modular	0.922	0.993
	(P=0.761)	(P=0.949)
Modular	1.190**	1.008
	(P=0.044)	(P=0.744)
Second inventor's knowledge depth ×		
Non-Modular	1.092	1.096**
	(P=0.502)	(P=0.015)
Modular	0.918	0.979
	(P=0.358)	(P=0.661)
Overlap in the knowledge scope of the first and second inventors ×		
Non-Modular	1.509*	1.126
	(P=0.056)	(P=0.303)
Modular	0.244*	0.975
	(P=0.060)	(P=0.590)
Third inventor's knowledge breadth ×		
Non-Modular	1.039	1.128***
	(P=0.571)	(P=0.000)
Modular	1.512***	0.896*
	(P=0.000)	(P=0.000)
Third inventor's knowledge depth ×		
Non-Modular	1.146	1.101***
	(P=0.585)	(P=0.000)
Modular	0.576***	1.095
	(P=0.009)	(P=0.356)
Fourth inventor's knowledge breadth ×		
Non-Modular	0.540***	0.901
	(P=0.000)	(P=0.160)
Modular	0.920***	1.132***
	(P=0.000)	(P=0.000)
Fourth inventor's knowledge depth ×		
Non-Modular	1.843***	0.953

		(P=0.001)	(P=0.420)
	Modular	1.288	0.908***
		(P=0.374)	(P=0.000)
Fifth inventor's knowledge breadth ×			
	Non-Modular	1.389	0.999
		(P=0.152)	(P=0.986)
	Modular	0.772***	0.925**
		(P=0.001)	(P=0.030)
Fifth inventor's knowledge depth ×			
	Non-Modular	0.690***	0.997
		(P=0.000)	(P=0.635)
	Modular	1.177***	1.088***
		(P=0.000)	(P=0.000)
Full set of controls		Yes	Yes
Technology-time fixed effects		Yes	Yes
Number of observations		4,110	16,905
R-squared		0.281	0.186

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table A7: Replicating Table 7 using OLS model

Estimation Model:	OLS	
DV:	**Novel breakthrough (1)**	**Economic breakthrough (2)**
Novel breakthroughs		0.048 (0.069)
First inventor's knowledge breadth ×		
modular × narrowly applied (MRI)	-0.001 (P=0.003)	0.003 (P=0.001)
modular × broadly applied (RFID)	-0.004 (P=0.000)	0.010 (P=0.000)
non-modular × narrowly applied (Stem Cells)	0.028 (P=0.000)	0.013 (P=0.002)
non-modular × broadly applied (Nanotubes)	0.004 (P=0.000)	0.023 (P=0.000)
First inventor's knowledge depth ×		
modular × narrowly applied (MRI)	0.000 (P=0.539)	0.005 (P=0.000)
modular × broadly applied (RFID)	0.001 (P=0.000)	-0.001 (P=0.039)
non-modular × narrowly applied (Stem Cells)	-0.010 (P=0.000)	-0.008 (P=0.000)
non-modular × broadly applied (Nanotubes)	-0.003 (P=0.000)	-0.009 (P=0.001)
Second inventor's knowledge breadth ×		
modular × narrowly applied (MRI)	-0.000 (P=0.785)	-0.002 (P=0.227)
modular × broadly applied (RFID)	0.008 (P=0.001)	0.001 (P=0.549)
non-modular × narrowly applied (Stem Cells)	-0.026 (P=0.000)	0.024 (P=0.002)
non-modular × broadly applied (Nanotubes)	0.006 (P=0.012)	0.001 (P=0.464)
Second inventor's knowledge depth ×		
modular × narrowly applied (MRI)	0.000 (P=0.289)	-0.000 (P=0.968)
modular × broadly applied (RFID)	-0.003 (P=0.004)	-0.010 (P=0.010)
non-modular × narrowly applied (Stem Cells)	0.013 (P=0.000)	-0.008 (P=0.014)
non-modular × broadly applied (Nanotubes)	-0.002 (P=0.132)	0.012 (P=0.010)
Overlap in the knowledge scope of the first and second inventors ×		
modular × narrowly applied (MRI)	-0.011 (P=0.002)	-0.003 (P=0.267)
modular × broadly applied (RFID)	-0.006 (P=0.016)	-0.021 (P=0.046)
non-modular × narrowly applied (Stem Cells)	0.048 (P=0.000)	-0.002 (P=0.281)
non-modular × broadly applied (Nanotubes)	0.007 (P=0.006)	-0.012 (P=0.056)
Full set of controls & Technology-time fixed effects	Yes	Yes
Number of observations	20,214	20,214
R-squared	0.319	0.145

Note: All models are OLS regressions with robust standard errors clustered at the technology level. P-values are reported in parentheses.

Table A8: replicating results - excluding RFID patents that do not have the term RFID in their title/abstract

Estimation Model:	**Logistic (Odds Ratios Reported)**
DV:	**Novel breakthrough**
First inventor's knowledge breadth ×	
modular × narrowly applied (MRI)	1.005 (P=0.343)
modular × broadly applied (RFID)	0.878 (P=0.000)
non-modular × narrowly applied (Stem Cells)	1.552 (P=0.000)
non-modular × broadly applied (Nanotubes)	1.393 (P=0.000)
First inventor's knowledge depth ×	
modular × narrowly applied (MRI)	0.926 (P=0.000)
modular × broadly applied (RFID)	1.494 (P=0.000)
non-modular × narrowly applied (Stem Cells)	0.744 (P=0.000)
non-modular × broadly applied (Nanotubes)	0.191 (P=0.000)
Second inventor's knowledge breadth ×	
modular × narrowly applied (MRI)	0.791 (P=0.000)
modular × broadly applied (RFID)	1.501 (P=0.000)
non-modular × narrowly applied (Stem Cells)	0.613 (P=0.000)
non-modular × broadly applied (Nanotubes)	1.565 (P=0.000)
Second inventor's knowledge depth ×	
modular × narrowly applied (MRI)	0.936 (P=0.368)
modular × broadly applied (RFID)	0.470 (P=0.000)
non-modular × narrowly applied (Stem Cells)	1.484 (P=0.000)
non-modular × broadly applied (Nanotubes)	0.338 (P=0.000)
Overlap in the knowledge scope of the first and second inventors ×	
modular × narrowly applied (MRI)	0.164 (P=0.000)
modular × broadly applied (RFID)	1.751 (P=0.201)
non-modular × narrowly applied (Stem Cells)	2.294 (P=0.000)
non-modular × broadly applied (Nanotubes)	1.189 (P=0.000)
Full set of controls & Technology-time fixed effects	Yes
Number of observations	3,675
Pseudo R-squared	0.274

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.

Table A9: Impact of team configurations on innovation outcomes using the full sample of patents

Estimation Model:	**Logistic (Odds Ratios Reported)**	
DV:	**Novel breakthrough (1)**	**Economic breakthrough (2)**
Novel breakthroughs		1.353 (0.009)
First inventor's knowledge breadth ×		
modular × narrowly applied (MRI)	1.000 (P=0.873)	1.032 (P=0.000)
modular × broadly applied (RFID)	0.861 (P=0.000)	1.179 (P=0.000)
non-modular × narrowly applied (Stem Cells)	1.558 (P=0.000)	1.161 (P=0.000)
non-modular × broadly applied (Nanotubes)	1.345 (P=0.000)	1.292 (P=0.000)
First inventor's knowledge depth ×		
modular × narrowly applied (MRI)	0.933 (P=0.000)	1.080 (P=0.000)
modular × broadly applied (RFID)	1.290 (P=0.000)	0.981 (P=0.039)
non-modular × narrowly applied (Stem Cells)	0.744 (P=0.000)	0.881 (P=0.000)
non-modular × broadly applied (Nanotubes)	0.188 (P=0.000)	0.884 (P=0.000)
Second inventor's knowledge breadth ×		
modular × narrowly applied (MRI)	0.793 (P=0.000)	0.974 (P=0.231)
modular × broadly applied (RFID)	1.385 (P=0.000)	0.976 (P=0.329)
non-modular × narrowly applied (Stem Cells)	0.599 (P=0.000)	1.094 (P=0.057)
non-modular × broadly applied (Nanotubes)	1.625 (P=0.000)	1.009 (P=0.463)
Second inventor's knowledge depth ×		
modular × narrowly applied (MRI)	0.954 (P=0.393)	1.002 (P=0.941)
modular × broadly applied (RFID)	0.677 (P=0.000)	0.904 (P=0.026)
non-modular × narrowly applied (Stem Cells)	1.551 (P=0.000)	1.002 (P=0.952)
non-modular × broadly applied (Nanotubes)	0.372 (P=0.000)	1.202 (P=0.000)
Overlap in the knowledge scope of the first and second inventors ×		
modular × narrowly applied (MRI)	0.159 (P=0.000)	0.972 (P=0.488)
modular × broadly applied (RFID)	0.925 (P=0.014)	0.846 (P=0.100)
non-modular × narrowly applied (Stem Cells)	2.300 (P=0.000)	1.356 (P=0.000)
non-modular × broadly applied (Nanotubes)	1.152 (P=0.000)	0.994 (P=0.845)
Full set of controls & Technology-time fixed effects	Yes	Yes
Number of observations	4,110	16,808
Pseudo R-squared	0.287	0.185

Note: All models are logistics regressions with robust standard errors clustered at the technology level. Reported estimates are odds rations. P-values are reported in parentheses.