Technovation 135 (2024) 103050

Contents lists available at ScienceDirect

Technovation

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

Quantifying advances from basic research to applied research in material science

Noriyuki Higashide [a,*], Yi Zhang [b], Kimitaka Asatani [a], Takahiro Miura [a], Ichiro Sakata [a]

[a] *Department of Technology Management for Innovation, School of Engineering, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, 113-0033, Tokyo, Japan*

[b] *Australian Artificial Intelligence Institute, Faculty of Engineering and Information Technology, University of Technology Sydney 15 Broadway, Ultimo, 2007, NSW, Australia*

A R T I C L E I N F O

Keywords:
Management of technology and innovation
Management science
Science and technology policy
Network science
Computational social science
Science of science and innovation

A B S T R A C T

Exploring the connection between basic science and its practical applications is critical to consider the social justification of the substantial governmental investments in science. However, there's a limited understanding of global research patterns and the dynamics of researcher collaboration from a basic-applied perspective because such studies have been mainly focused only on the biomedical field. Here, the main goal is to propose an indicator to quantify the degree of basic-applied research in academic papers. Using the indicator, we uncover how material science has advanced from basic to applied research, based on the international trends of the indicators and the affiliations of the scientists involved. We develop a methodology that indexes levels of advancement from basic research to applied research based on large-scale text data. The continuous scores assigned to each paper are derived from a vector space embedding technical terms from a broad network data. These scores align with experts' views in material science. This methodology enables us to monitor international trends that China has significantly advanced into applied research, as well as Chinese applied scientists increasingly associating with their domestic institutions. As science and technology implication, our methodology extends the boundary of assessing scientific research on its proximity to real-world applications and provides a tangible measure for funding agencies managing to fund or design research environments.

1. Introduction

Governments invest in scientific research to expand our scientific understanding, ignite innovation, drive economic progress, and enhance societal welfare. This activity seems to be motivated by Linear Model, which is commonly understood to progress from basic research to applied research and eventually to real-world applications (Balconi et al., 2010; Bush, 2020). In biology, for instance, once basic research at the cellular or genetic level uncovers promising leads for particular diseases, advancing toward drug development, and then, through clinical research and trials, novel drugs hit the shelves. In electronics, after quantum mechanics theory was largely completed, band theory allowed control of electrons at a quantum level, which led to the age of semiconductor engineering, and today, semiconductor devices are indispensable for maintaining civilization. Considering that these processes happen in various fields, it becomes clear that investing in basic science is utterly essential for our society to better drive applied research and commercialization.

The motivation for political funding for science and technology has been driven by the narrative that investing boosts national security, creates jobs, improves public health, and enhances the quality of life, all contributing to the nation's wealth (Macilwain, 2010). This approach stems from Vannevar Bush's post-war recommendations (Bush, 2020). Investment in basic research was primarily touted for such benefits, yet the narrative was largely based on anecdotes or weak rationals lacking quantitative or objective evidence (Macilwain, 2010; Lane, 2010). In response to such trends, approaches that utilize multiple data to assess outcomes of funding in science and technology were proposed in 2010 exemplified by STAR METRICS (Lane and Bertuzzi, 2011). Lane and Bertuzzi (2011) suggests the creation of a data infrastructure that compiles scientific outcomes in as open as possible. It argues for the necessity to quantitatively evaluate outputs like grant proposals, awards, publications, and patent connectivity, in order to conduct cost-benefit analysis for future investment strategies.

Since then, the economic, technological, and academic benefits of funding in science and technology have been evaluated using data.

* Corresponding author.
E-mail address: n.higashide@ipr-ctr.t.u-tokyo.ac.jp (N. Higashide).

https://doi.org/10.1016/j.technovation.2024.103050
Received 14 October 2023; Received in revised form 3 May 2024; Accepted 23 May 2024
Available online 9 June 2024

Weinberg et al. (2014) showed that research and development expenditures by the federal government have supported scientific research activities, leading to the employment of a diverse range of talents, including students, postdocs, and research staff, and to the consumption of a variety of goods and services purchased locally, nationally, and across the country. Some studies have examined the relationship between research and development spending and economic indicators (Coccia, 2009, 2018b). Technological benefits have been mainly assessed by measuring academic papers' citations from patents (Jefferson et al., 2018; Ahmadpoor and Jones, 2017). Li et al. (2017) revealed that 10% of NIH-funded academic research directly contributes to patent development, showcasing a tangible return on public investment in science. Patents in the U.S. increasingly depend upon federally supported research and it has reached nearly 30% in 2019 (Fleming et al., 2019). For academic benefits, funded research projects appear in more prestigious journals and attain more citations than non-funded ones (Yan et al., 2018; Álvarez Bornstein and Bordons, 2021; Roshani et al., 2021; Mosleh et al., 2022; Coccia and Roshani, 2023, 2024). As shown here, data utilization to inform funding decisions has become increasingly important.

In recent years, Science, Technology and Innovation (STI) policies have been at a turning point, requiring a broader perspective on investments in scientific research. Amid global challenges such as pandemics, climate change, and security crises, the OECD has launched initiatives urging countries to shift their STI policies to address these crises (OECD, 2023). Collaborating globally to drive innovation is essential to tackle these worldwide challenges, making research investment crucial. During this pivotal time in STI policy, understanding how scientific research in academia can lead to innovation in the real world is becoming essential for funding agencies.

Yet, funding agencies increasingly struggle to accurately assess the values of complex and specialized science. Traditionally, research topics have been evaluated through expert discussions, but such qualitative assessments are limited by the experts' knowledge and quality. The number of academic papers is growing exponentially, potentially overlooking signals of emerging science and technology trends or rising new fields worldwide. To address this issue approaches combining large-scale data and analytical methods have been proposed to grasp global trends in cutting-edge scientific and technological innovations (Athey, 2017; Baas et al., 2020; Liu et al., 2023), enabling funding agencies to obtain antenna-like capability. They can quantitatively evaluate science using various indicators such as trends in research papers and patents, citations and novelty as a measure of impact, and researchers' attributes and performance.

Research portfolios should also be considered based on whether research projects lean towards basic or applied research (Dudley, 2013). Understanding the distance to practical application when aiming for a certain direction of innovation can support funding agencies' decisions. This concept has traditionally been thought of in terms of Technology Readiness Levels devised by NASA, based on qualitative discussions and limited by the fields it applies to and the participating experts' knowledge (Nakamura et al., 2013). Designing a data-driven indicator based on knowledge across various is critical for the future of funding for science, technology and innovation.

The goal of this study is to propose a method that uses large-scale data to calculate an indicator determining where academic papers fall on the spectrum between basic and applied research. Further, the other goal is to use this indicator to uncover valuable information that can support data-driven decision-making. Investigating how scientific advance follows the Linear Model and exploring scientists' collaborations are fundamental questions in science and technology, as we review in Section 2. Therefore, this indicator is utilized to gain new insights into the evolution of scientific fields and the collaboration among scientists.

We choose nanocarbon materials as a case study because the field has advanced from basic research through applied studies and finally to commercialization in the past 50 years, suggesting a perfect case to evaluate a novel quantitative indicator (Li et al., 2014; He et al., 2018; Soliman et al., 2020). When using metrics to assess the development of fields, the study expands the scope to photocatalysts and batteries, confirming the versatility of the metrics within the broader context of material science. Regarding collaboration among researchers, the focus is particularly on those from Japan, the US, and China in the nanocarbon materials field, where there is a high volume of publications.

2. Literature review

We will review the definitions and relationships between basic and applied research, as well as their identification methods, which are closely linked to the main topic. Assuming the advance from basic to applied research, we also discuss related topics such as the science-technology linkage and the evolution of scientific disciplines. Furthermore, understanding the evolution of science includes examining the affiliations of scientists as a background factor, which relates to and will be reviewed under the topic of collaboration in science.

To argue basic and applied research, we use Pasteur's Quadrant which Donald Stokes challenged the Linear Model's simple dichotomy of basic versus applied research and proposed a two-axis classification (Stokes, 2011). The one axis questions if research considers practical use, and the other, if it aims for fundamental understanding (Table 1). Based on these two axes, the classification includes: 1) Pure basic research, not aimed at practical use but driven by curiosity and seeking fundamental understanding, exemplified by Neils Bohr, 2) Pure applied research focused on practical use symbolized by Tomas Edison, 3) Use-inspired basic research, motivated by practical goals and contributing to fundamental understanding of phenomena, fits Louis Pasteur, 4) unnamed group, not aim for basic understanding neither practical use (Dudley, 2013).

Here, basic research is the type of research that generally falls into 1) Pure basic research within Pasteur's Quadrant. Their objective is to understand pure scientific phenomena or fundamental properties. For example, studies of the fundamental mechanisms of DNA and examining unknown physical properties of new nanomaterials. However, even in basic research, it's common to secure grants with practical applications in mind, often aligning with category 3) Use-inspired basic research. For instance, understanding nanomaterials could benefit next-generation electronics, like integrated circuits or solar cells. Also, understanding how the brain forms new neural connections might suggest novel treatments for brain injuries.

Applied research is executed mainly to explore practical solutions to address specific situations or issues. This often fits category 2) Pure applied research. Examples of applied research include developing new clinical treatments for certain cancers or improving the performance of silicon solar cells. Yet compared to Edison's time, technology has become more specialized and interdisciplinary, requiring foundational scientific understanding and expertise across multiple fields. In some fields, part of category 3) might also be considered as applied research.

Basic research and applied research influence each other. It's commonly believed today that research and development proceed through a sequence of basic research, application, and productization, known as the Linear Model (Balconi et al., 2010; Bush, 2020). General scientific knowledge derived from pure basic research can lead to innovative technologies in the next generation represented by the flow

Table 1
Pasteur's quadrant (Stokes, 2011).

	not considers practical use	considers practical use
aims fundamental understanding	1) Pure basic research	3) Use-inspired basic research
not aims fundamental understanding	N/A	2) Pure applied research

that quantum mechanics guided semiconductor transistors. However, the advance of science and technology is not limited to this pattern. Products released through applied research for productization can enable subsequent basic research, expanding the frontier of knowledge (McKelvey, 1985). New technologies from pure applied research not aimed at fundamental understanding can shape the next generation's pure basic research. When the steam engine was created, there was no explanation or theory for its operation; the theory of thermodynamics was established later. Additionally, understanding fundamental phenomena requires observational technologies; technologies like telescopes and clocks were essential for verification of laws of motion explaining celestial movements.

In fields like advanced materials and biopharmaceuticals, the Linear Model is more prevalent because these technologies must ensure safety and industrial stability before reaching the stage of societal implementation at scale. Indeed, many basic researchers in such fields are required to demonstrate potential applications of their research when seeking grants. For example, they are the impact of genetic traits on diseases or how quantum-level behavior of electrons or photons could become key technologies for next-generation computer (Dudley, 2013). In this study focused on advanced materials, we aim to identify where academic papers lie on the advance from basic to applied research.

The Linear Model assumption has been the mainstream trend from the post-war period to today (Pielke Jr, 2012; Bush, 2020). After the Second World War, following the successful military application of basic research, Vannevar Bush proposed to President Roosevelt how to conduct scientific research as a nation (Bush, 2020). Under the assumption of linear advance from basic to applied, budget allocation mechanisms for science and technology have been designed not only in the U.S. but also in many other countries. The assumption of progress from basic to applied research and then to commercialization is also evident in studies that explore how fundamental university research diffuses into corporate innovation (Cao et al., 2023) and the impact of technology transfer in big science (Scarrà and Piccaluga, 2022).

Common techniques used to identify the position of a paper in the spectrum from basic to applied research can be classified into two categories: those using textual information and those using the structure of citation networks. Our study aims to enhance the text-based approach. Originally, biomedical journals were manually classified into four research levels to understand the structure of biomedical literature (Narin et al., 1976). In 2004, the classification of research levels for roughly 3000 journals was accomplished by text processing. By calculating the percentage of article titles containing a particular group of words identified as "basic" or "clinical," classification is consistent with Narin's proposed levels (Lewison and Paraje, 2004). A machine learning model that used textual information has been presented that classifies individual articles into four research levels using their titles, abstract words, and cited references (Boyack et al., 2014). In a study of machine learning, a unique five-stage category ranging from basic to applied research distinct from Research levels was devised and allowed for high-accuracy classification (Surkis et al., 2016). A machine learning system was also developed to identify if an article is likely to be referenced in future clinical trials or guidelines (Hutchins et al., 2019). Predefined technical terms assigned to each biomedical article, MeSH terms, were used to build a technique that creates a basic-to-applied map (Weber, 2013). In 2019, a method was developed to assign a continuous research level to over a million articles by analyzing the co-occurrence network of these technical terms (Ke, 2019). This value is defined as Level Score, which is also well aligned with the research level proposed by Narin et al. (1976) and allows for more quantitative analysis of the spectrum from basic research to applied research. However, the technique is only used in the medical field as it necessitates using human-maintained technical terms linked to the articles, which are MeSH terms.

The impact of scientific papers on technical applications is also measured by how patents cite them (Narin et al., 1997; Ahmadpoor and Jones, 2017; Ke, 2020a, 2020b; Manjunath et al., 2021). These are the methodology based on the paper-patent citation network. Some work has expanded on the paper-patent relationship to government grants. They investigate how government grant-supported research resulted in publications, patents, and products. For instance, the majority of the impact of NIH-granted research on patents is indirect (Li et al., 2017). Also, those applying for patents rely on research supported by federal funds as a knowledge base (Fleming et al., 2019). According to these reports, both basic science and applied science have an equal impact on innovation in drug discovery (Du et al., 2019; Ke, 2020a).

The interconnection between science and technology is crucial for fostering innovation, particularly evident in fields like nanotechnology, where scientific knowledge across physics, chemistry, and material science (Islam and Miyazaki, 2009, 2010). Mansfield (1991) argues if it were not for the contribution of academic research, 10% of new products and manufacturing methods would have been significantly delayed in their introduction. The significance of science in driving economic growth through innovation is acknowledged (Narin et al., 1997), with patent citation analysis serving as a method to gauge the science-technology connection (Albert et al., 1991; Narin and Breitzman, 1995; Michel and Bettels, 2001). Recent studies utilize network analysis to identify emerging research fronts by analyzing academic paper citation networks, revealing gaps between science and technological applications, like in solar cell development (Shibata et al., 2010, 2011). Semantic similarity and publication time lag between papers and patents are also evaluated to understand their interplay better (Ogawa and Kajikawa, 2015; Yang et al., 2023). Additionally, the use of computational approaches for monitoring emerging technologies (Kyebambe et al., 2017; Ogawa et al., 2018; Li et al., 2019) and assessing "start-up readiness" in the biopharmaceutical sector (Goji et al., 2020) demonstrates the growing reliance on quantitative analysis to explore the dynamics of science-technology linkages, which are becoming more apparent across various fields thanks to enhanced network analysis techniques and data availability (Ahmadpoor and Jones, 2017; Ke, 2020a, 2020b; Manjunath et al., 2021).

Mario Coccia's studies explore the evolution of scientific disciplines, highlighting key dynamics and influences. His 2018 research indicates that a few dominant disciplines often drive field evolution (Coccia, 2018a), and his 2020 study on a theoretical framework of scientific evolution details the necessary for societal progress made changes (Coccia, 2020). "Technological parasitism" explains how technologies coevolve through interdependent relationships (Coccia and Watts, 2020). In addition to theories and conceptual frameworks about the sources and failures of innovation (Coccia, 2017, 2023), Boyack et al. (2005) have presented a new map depicting the overall structure of science based on scholarly articles. Such efforts reflect ongoing investigations into the evolution of science through generalizable theories derived from case studies and empirical studies. However, examining the evolution of specific fields through the lens of the basic-to-applied science continuum using large-scale data has yet to be conducted.

Understanding the patterns of international collaborations is also an important piece for a comprehensive picture of the evolution of science. Thus, it has been discussed, with a global increase and a rising presence of countries like China, moving beyond a Western-centric view (Luukkonen et al., 1992; Adams, 2012). Historically international collaborations are more common in basic research fields and less so in applied research, the gap between these fields has been narrowing (Coccia and Wang, 2016; Coccia and Bozeman, 2016; Bozeman and Youtie, 2016). This may suggest that international boundaries in many areas are becoming increasingly blurred. Pan et al. (2012) demonstrated that the strength of collaborative research between cities decreases as the distance between them increases and that a country's total research impact linearly grows in proportion to its research and development funding.

There have been studies focusing on sociological approaches, like theoretical frameworks, to explain the dynamics and outcomes of research collaboration among scientists. Bozeman and Corley (2004) have examined how scientists at academic research centers in the U.S. utilize "Scientific and Technical Human Capital" through research collaboration. It has revealed that tenure, larger grants, and gender significantly influence collaboration patterns and strategies. Collaborations have been categorized into knowledge-focused and property-focused, proposing a new analytical framework and offering improvements for future research on collaboration dynamics and outcomes (Bozeman et al., 2013). Bozeman et al. (2016) highlights factors to influence the quality of collaborative efforts, by analyzing interview contents to identify good and bad collaboration dynamics with 60 U.S. academic researchers.

On the other side, there also have been lots of mathematical approach, network analysis and agent model, to understand collaborations in science. The study has explored the structure and dynamics of scientific collaboration networks, highlighting the "small world" characteristics of such networks where scientists are closely connected, often separated by only a few steps (Newman, 2001). The distribution of coauthorships tend to follow a power-law (Newman, 2004). Factors like team size, newcomer ratios, and repeat collaborations shape network structure and performance (Guimera et al., 2005). The studies also demonstrate that an author's centrality in a network correlates positively with the impact of their research (Uddin et al., 2013) and that disciplines evolve through social interactions within these networks, influencing the broader scientific landscape (Sun et al., 2013). Each piece underscores the importance of understanding collaborative processes to optimize scientific innovation.

3. Methods

3.1. Research setting

As we mentioned in Section 1, the research questions for this study are.

- How can basic and applied research be accurately quantified?
- How does the research focus of a country shift from basic to applied research over time?
- How does scientific collaboration change during the transition from basic to applied science?

These questions aim to explore the effectiveness of quantitative measures in understanding and visualizing the dynamics of scientific research and collaboration.

As a case study, material science is selected for two main reasons to explore these questions. Firstly, it's a rapidly evolving field spanning basic to applied research, which allows us to verify if the method can accurately capture this growth. Secondly, it's a very large field with many researchers and publications, and diverse including areas like batteries, solar cells, biosensors, and composite materials (Islam and Miyazaki, 2009, 2010). Nanocarbons, in particular, are an appropriate subfield to analyze because of the clearer advancing stream from understanding basic scientific properties to exploring potential applications.

Nanocarbon materials, encompassing fullerenes, carbon nanotubes, and graphene, offer promising avenues for future structural materials, electronic gadgets, and electrochemical substances due to their lightweight nature, unparalleled thermal and electrical conductance, and specific electrochemical traits (Li et al., 2014). Initially, the focus on basic research has arisen, and these materials have gradually become the subject of applied studies, showcasing a rise in interdisciplinary connections (Maruyama, 2002; Maruyama et al., 2002; He et al., 2018). Recent research endeavors are actively pivoting towards real-world applications like biosensors, electrocatalysts, nano-scale light sources, cellular biotechnology (Begum et al., 2011; Jiao et al., 2015; Justino et al., 2017; Higashide et al., 2017; Chu et al., 2021). Commercialization efforts have notably accelerated in the past decade, exemplified by a surge in patent activities and the launch of innovative products (Higashide et al., 2023). A particular product, comprising ultra-fine carbon particles, serves as an additive to improve thermal and electrical properties in various materials, such as cement (Soliman et al., 2020). The field of nanocarbon materials is transitioning from a research-centric phase to one of commercial viability. Given that these materials are carbon-based and exhibit unique physical properties, they are increasingly seen as sustainable substitutes for traditional materials in a world moving towards decarbonization.

The assumptions in this study can be summarized as follows.

- The majority of science and technology advances from basic research to applied research and finally to commercialization.
- Government funding for science and technology is often based on this pattern.
- Research on nanocarbon materials has progressed globally according to this pattern.
- For funding agencies, the need to use large-scale data to gain an overview of national development and scientific collaboration is increasing.

3.2. Sample and data

We prepared detailed research paper data from the Scopus Custom Dataset by Elsevier, covering all the documents it has listed from 1970 to 2020. Scopus is the largest database for summaries and citations of reviewed scholarly articles. With 73 million papers and 1.2 billion citations in almost every research area, it is larger than other similar databases, and hence researchers widely use Scopus data to study bibliometric and citation patterns (Asatani et al., 2018, 2023; Baas et al., 2020; Miura et al., 2021).

We first built a computational environment to handle the large Scopus dataset using Python and then extracted papers that contained both "carbon nanotube" and "luminescence" in their titles, keywords, or abstracts. 2,500 papers are obtained but the number is small for analysis. Consequently, we expanded our initial data to include papers directly connected to the citation network, which means including both citing and cited papers, totaling 81,634 papers (Fig. 1a). This approach allowed us to broadly cover adjacent papers related to the luminescence of carbon nanotubes, namely those about nanocarbon materials and their functional properties. The paper covers the period from January 1970 to December 2020, providing information sufficient to observe the advances in this field over 50 years.

3.3. Measures and variables

To address the research questions, it is necessary to define and measure specific indicators for each. First, we need a method and variables to quantify the extent of research from basic to applied, which will be discussed in the next subsection. For the second and third questions, it is essential to assign nationalities to research papers. Apart from the level of basic to applied research, the impact of papers will be assessed using citation counts as reference information.

In our Scopus dataset, the organization that each author belonged to at the time of publication is recorded for each paper, and each organization is linked to their nationality. Some papers are written solely by authors from within the same country, while others are collaborative efforts involving authors from organizations across multiple countries. The nationality of each scientist was defined as the nationality of the organization when they wrote their first paper in their career (Sugimoto et al., 2017).

To estimate the impact of each paper, we used the number of citations it received within 10 years after publication C_{10} (Sinatra et al., 2016). We then normalized C_{10} by dividing it by the mean citations in the paper's field and publication year in order to calibrate for variations in publication frequency and impact across different fields (Radicchi

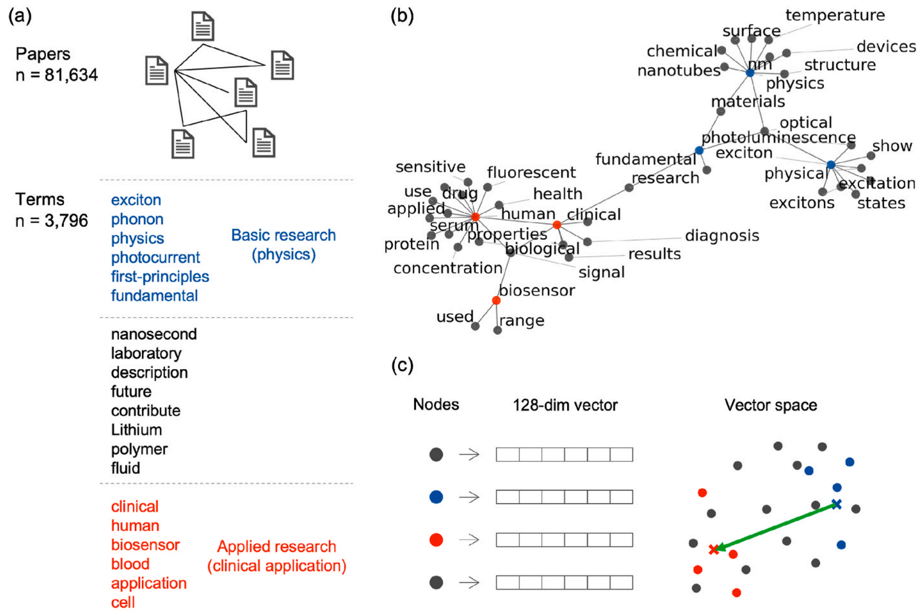

Fig. 1. Illustrating term-based methodology to quantify the degree of scientific advance. (a) Papers related to nanocarbon materials and their functional properties. From these, specialized terms are extracted, with those in blue associated with basic research, and those in red linked to applied research. (b) The term co-occurrence network. The edges are weighted by co-occurrence frequency, and the nodes are 3,796, equivalent to the number of extracted terms. The figure shows the part of the network whose edge weight is high. Blue and red nodes represent terms relevant to basic and applied research, respectively. (c) The concept of vectorizing the co-occurrence network. Each node is converted into a 128-dimensional vector using LINE. In the vector space, an imaginary vector is defined from the centroid of the basic nodes to that of the applied nodes shown as the green vector from the blue cross mark to the red cross mark. The cosine similarity between this vector and the others offers a quantified value, Topic-Aware Level Score, reflecting the advance from basic research to applied research.

et al., 2008). The fields are identified using the codes assigned to the Scopus dataset. We used a log-transformed value with a base of 10 since normalized C_{10} follows a log-normal distribution.

3.4. Models and data analysis procedure

We developed the methodology to identify levels from basic to applied research, drawing upon a former study (Ke, 2019). This model uses a large corpus of academic papers and a specialized terminology dictionary as inputs. It obtains vector representations of co-occurrence networks of terms, then the extent of research from basic to applied is calculated.

The previous study required a dataset of specialized terms linked to each paper, but in the targeted nanocarbon field, no such pre-defined data exists. We used the Latent Dirichlet Allocation (LDA) topic model to automatically extract technical terms (Blei et al., 2003). LDA topic model is a widely-used unsupervised method, meaning we can explore a holistic view of latent topics without much attention and computational cost. We began with assembling a corpus from our nanocarbon paper collection by combining the title, abstract, and keywords for each paper, and then we removed any symbols and stop words. Next, We defined eight topics and obtained the 500 most frequent terms for each. After eliminating duplicates, 3,796 unique technical terms were left.

Some of the terms are needed to be categorized as either basic or applied research, but it's not practical to examine each term individually. In the prior research we referred to, terms belonging to the "cell" or "animal" categories in MeSH dataset are labeled as basic research, while words in the "human" category are labeled as applied research (Ke, 2019). Inspired by that approach, we labeled terms utilizing the eight topics. The topic *A* is represented by technical terms such as "quantum", "phonon" and "physics", indicating fundamental physical properties of nanocarbon materials. The another topic *B* is characterized by "biosensor" or "clinical", which represents clinical applications. Therefore, 437 terms associated with topic *A* are labeled as basic research, and 437 terms from topic *B* are applied research. A sample of these terms and their labels is presented in Fig. 1a. Note that we found also that the topic *A* and topic *B* were the farthest apart in the two-dimensional space of principle component analysis.

We then weaved the terms into a co-occurrence network; Each term forms a node, and the edges are weighted to reflect the number of co-occurrence (Fig. 1b). This exercise allows us to define semantic distances between the terms. Surrounding the basic terms highlighted in blue are words such as "temperature", "optical", and "states", which denote the physical characteristics of materials. Around the red nodes, terms such as "diagnosis", "applied", and "human" illustrate applications of the materials.

Obtaining vector representations of nodes is called network embedding. Embedding a co-occurrence network helps computers understand terms' meanings better with lesser computational power and a broader range of data processing. Large-scale Information Network Embedding (LINE) is one of the embedding techniques. We employed LINE because it is more computationally efficient for large-scale networks and able to preserve higher-order proximity compared to other method such as node2vec (Tang et al., 2015; Grover and Leskovec, 2016).

By using LINE, we obtained 128-dimensional vectors for each of the 3,796 terms (Fig. 1c). Note that 128 is the default parameter setting of the LINE package. These vectors numerically reflect the relative semantic differences and therefore, the center of gravity in the vector space of terms labeled "basic" represents the conceptual center of basic research. Considering this, we next defined the vector from the centroid of terms labeled as basic research to that of terms labeled as applied research. The green vector shown in Fig. 1c, Translational Axis, serves as our compass - pointing from the realm of fundamental understanding to practical application. By calculating the cosine similarity of each term with the translational axis, the degree of advance towards application was quantified.

We named the value Topic-Aware Level Score. This is a continuous measure indicating the advance from basic to applied research presented in Section 2. Note that the similarly calculated value was named Level

Score in prior research (Ke, 2019). We have adopted the different name due to this method drawn on the terms extracted using topic model. A unique contribution is its ability to calculate scores in areas without a specific terminology dictionary.

To validate the calculated Topic-Aware Level Scores, we performed the expert assessment with two researchers who were asked to assign scores from basic to applied on a five-level scale. Text-based data analysis results are commonly evaluated by experts to ensure the validity of the findings (Zhang et al., 2014, 2016, 2018). To ensure a balanced selection across a range of the scores, papers were randomly chosen: those with scores below the 20th percentile, those with scores above the 80th percentile, and those with scores falling between these two percentiles. 10 papers were selected from each and 1 paper was excluded due to incomplete information. The interviews were conducted through online surveys. Based on the papers' titles, journals, publication years, and keywords, the experts were asked to classify them into five research levels: "basic", "somewhat basic", "intermediate", "somewhat applied", and "applied". Two experts were researchers in the field of nanocarbon materials at the School of Engineering, The University of Tokyo. Expert *A* is an assistant professor and Expert *B* is a professor.

Aiming for the second goal of understanding the development of science by country using metrics, we calculated the Topic-Aware Level Score for papers and plotted the dependency on publication year for each country. To ensure the applicability of this method beyond just one field, we focused on two subfields of material science: photocatalysts and batteries (Fig. 5d–i). Both have growth patterns similar to nanocarbon that Japan and the U.S. initially led in these areas, but China later dominated in their applied research (Ibhadon and Fitzpatrick, 2013). For photocatalysts, we gathered papers that included the following keywords; "photocatalyst", "photocatalysis", "photocatalytic materials", "titanium dioxide", and "zinc oxide" in their title or abstract. This resulted in a collection of 128,386 papers. For batteries, the query terms were "battery material", "lithium ion battery", "anode material", "cathode material", "solid state battery" and "sodium ion battery", yielding a collection of 52,113 papers. Publication periods range 50 years from January 1970 to December 2020 for each collection. In both domains, we created a corpus and applied the LDA topic model for term extraction. Subsequently, terms were labeled under expert knowledge, and Topic-Aware Level Score was calculated for each paper using network vectors.

To effectively visualize trends in scientific collaboration and understand the underlying factors differentiating the research focus between countries, we consider the transition of China into applied research compared to the continued emphasis on basic research in the U.S. and Japan in nanocarbon materials as the case. This shift could be attributed not just to China's rapid economic growth and significant government investment, but to unique aspects of the research environment. By applying Topic-Aware Level Score and data-centric analysis, we can plot and examine these national trends over time, providing clear visual insights into the dynamics of global scientific collaboration and development strategies.

4. Results

On examining Topic-Aware Level Score of individual terms, an intriguing pattern emerges; terms associated with basic research tend to score lower, while those tied to applied research bear higher scores (Fig. 2a). For instance, in the area of −0.5, we can see terms such as "phonon" that have a basic label and obviously physics-related term "atomistic" that isn't labeled. By contrast, around 0.5, we notice the term "biomarker" that isn't labeled but leans towards application. By averaging the scores of terms in a paper, we offer a quantitative measure of the extent to which it ranges from basic to applied research (Fig, 2b). Papers on the basic side of the distribution curve are ones like those in Physical Review B discussing exciton scattering. On the applied research end, there are papers on immunosensors using carbon nanotubes.

We find a pattern that physics journals tend to have lower Topic-Aware Level Score, general and chemistry-focused journals are nestled in the middle, and journals dedicated to biosensors mark higher scores (Fig. 3). We see a clear flow from low to high with the order of physics-focused, chemistry-focused, and biosensor-focused (see also Supplementary Fig. S1). This suggests that the Topic-Aware Level Score seems to accurately reflect the advance from basic research in physics to applied research in biotechnology. In the world of carbon nanotubes, it might just be a solid tool for tracking the advance of research from theory to application.

To further confirm the consistency of Topic-Aware Level Score, we compare them with assessments performed by two experts in the nanocarbon field. The qualitative classification varied between the two experts. Expert *A* categorized 10 papers of the highest number as "basic" (Fig. 4a). In contrast, Expert *B* classified 17 papers as somewhat applied (Fig. 4b). These reflect the variation in criteria for qualitatively assessing the distinction between basic and applied research. There is a positive correlation between the Topic-Aware Level Score and the five research levels. The correlation coefficients, higher for Expert *A* of 0.73 than for Expert *B* of 0.37, indicate that Expert *A*'s judgment criteria are closer to our Topic-Aware Level Score. Our framework is well aligned with the qualitative assessment by expert interview.

To characterize the trend of the nanotube innovation, we compared publication count and Topic-Aware Level Score across countries. The volume of publications has risen exponentially, aligning with general academic trends, particularly in the nanocarbon field (Fig. 5a). Fig. 5b illustrates a shift from "basic" to "applied" research across all countries. Japan and the U.S. have consistently shown lower Topic-Aware Level Score than the global average, with publication peaks around 2010 followed by declines, suggesting a strong foundation in basic research from earlier decades. Conversely, China, initially trailing, surged in publication count in the 2000s, reflecting a shift towards applied research. South Korea and India align with the global average trend; however, India's publications are increasing, unlike Korea's. This

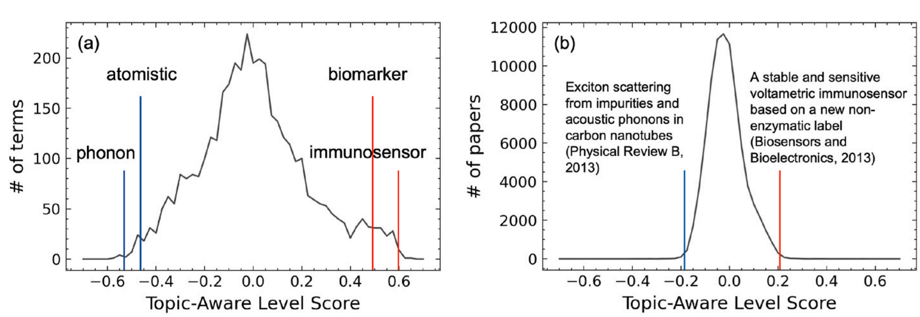

Fig. 2. (a) The distribution of Topic-Aware Level Score for each term is displayed. Lower values are assigned to words leaning towards basic research, while higher values go to those favoring applied research. (b) The distribution of Topic-Aware Level Score for research papers is obtained by averaging the scores of terms contained in a paper.

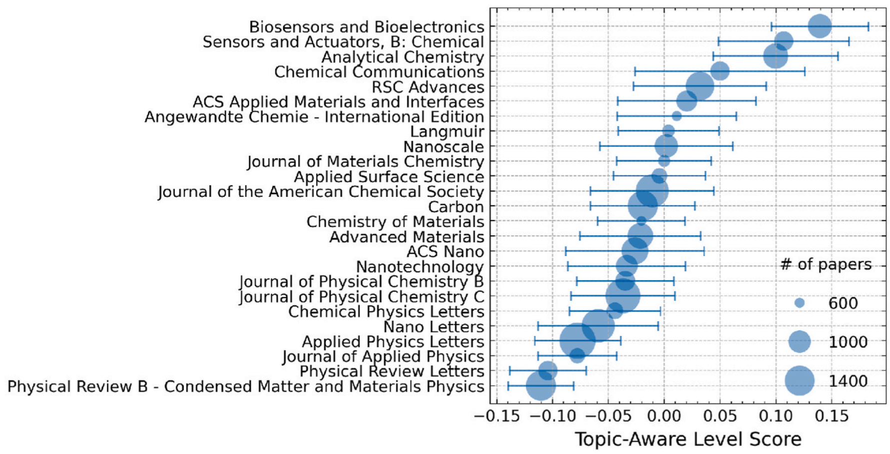

Fig. 3. The mean of Topic-Aware Level Score for each journal calculated from the collection of nanocarbon papers. The size of the circle and the length of the bars indicate the number of papers and the standard deviation of the score respectively. Journals with more than 600 publications are shown.

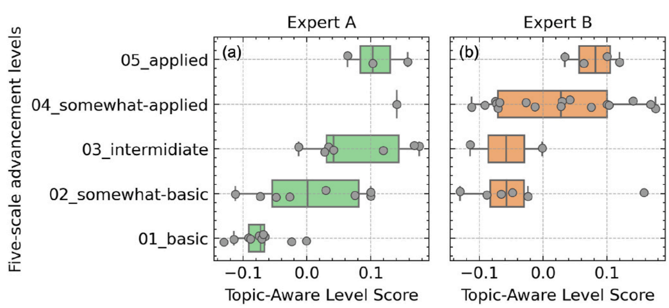

Fig. 4. Expert assessment based on title, keywords, and publisher of each paper randomly selected from nanocarbon paper datasets. The vertical axis shows qualitatively assessed five-scale advancement levels. In the box plots, the average values of Topic-Aware Level Score for each category are shown and each dot represents paper. We assign scores of −0.5, −0.25, 0.0, 0.25, 0.5 to the five levels, correlation coefficients with the horizontal axis are 0.73 and 0.37 for the expert *A* and *B*, respectively which suggests Topic-Aware Level Score is consistent with the qualitative estimation by experts.

pattern fits a transition from foundational research in the U.S. and Japan to applied research in China, verified by the expert hearing.

China exhibits a unique trend, with a significant boost in both publication numbers and Topic-Aware Level Score since the 2000s, positioning it at the top globally. This trend is also maintained in other materials science fields. For photocatalysts and batteries, China's Topic-Aware Level Score remains higher than other countries (Fig. 5c–f). The number of publications reached top level in the past 20 years since 2000 (Fig. 5c–e). Both photocatalysts and batteries are fields where commercialization is expected, and applied research is becoming active in China (Islam and Miyazaki, 2010; Ibhadon and Fitzpatrick, 2013). We have succeeded in quantitatively understanding the macro trends widely recognized in materials science and engineering, which indicates that Topic-Aware Level Score is a valid framework for capturing macro innovation trends.

The year 2010 may have been a turning point in nanotube research. That year, China became the leader in the number of papers. At the same time, the U.S. and Japan, which had been in the first and second places, peaked and have been decreasing since then (Fig. 5a). As for Topic-Aware Level Score, China is increasing at a greater rate than the U.S. and Japan (Fig. 5b). This reflects that China has newly entered applied research, while the U.S. and Japan have remained in basic research. What are the factors behind this difference? For one thing, the research environment in China, which has been successful in applied research, would be different from that of the U.S. and Japan. Rather than the generally stated differentiating factors, such as China's dramatic economic growth and the government's large-scale investment, here we

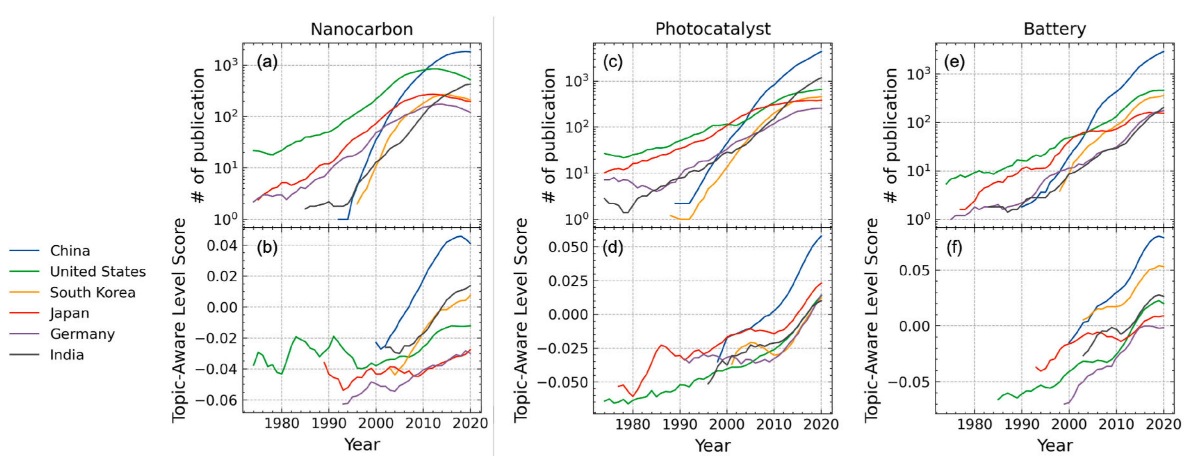

Fig. 5. Trends in the number of publications, Topic-Aware Level Score for each country for (a,b) papers on nanocarbon (n = 81,634), (c,d) papers on photocatalysts (n = 128,386), and (e,f) papers on batteries (n = 52,113). For each year, the mean of Topic-Aware Level Score of papers is shown. Years with fewer than ten papers are excluded from the data points. The nationality of the paper is the nationality of the organization in which it is published, and the top six countries with the largest number of nanocarbon papers are selected. Note that all graphs are expressed as 5-year moving averages to reduce noise.

analyze this difference from a data-centric perspective focused on the authors.

To understand the characteristics of scientists engaged in applied research, we examined the trend of the proportion of applied scientists belonging to organizations in their own country. Applied scientists are authors of papers published since 2010 whose Topic-Aware Level Score is in the top 10%. While Chinese scientists are increasingly affiliated with domestic organizations, scientists in Japan and the U.S. are on a downward trend, becoming affiliated with organizations abroad (Fig. 6). As shown by the solid blue line, Chinese applied scientists of around 10,000 belong to domestic organizations at a rate of more than 90%, which is higher than other scientists shown by the dotted line. Applied scientists in the U.S. and Japan are much fewer than those in China and have a lower proportion of affiliation with domestic organizations than general researchers. The peculiarity of Chinese applied researchers may be one of the factors explaining the advancement of applied research in China.

For comparison with another group, we performed a similar analysis targeting impactful scientists who have written papers with top-10% citation counts. Chinese impactful scientists have a roughly 10-point lower proportion of home-country affiliation than applied scientists, which indicates they are not as localized only in China as applied scientists are (Supplementary Fig. S2).

5. Discussion

The primary results of the analyses for the three research questions were summarized in Table 2, highlighting the contributions they make to existing knowledge.

Firstly, by using the quantification indicator Topic-Aware Level Score (TALS) to identify basic and applied research, the results could be interpreted using the theoretical framework, Pasteur's Quadrant (Table 1). Papers with low TALS, predominantly from physics journals (Fig. 3), are generally categorized as 1) Pure basic research. For instance, elucidating the energy band structure of a newly discovered material would fall under this category. Papers near a TALS of zero likely appear to 3) Use-inspired basic research because the primary focus is on chemistry journals, aiming to understand the fundamental phenomena of nanotube chemical reactions while also considering their potential applications. Indeed, many recent studies likely fall into category 3) regardless of whether they are oriented toward physics, chemistry, or medicine, as most research projects are designed with potential future applications. Papers with high TALS, which often relate to biosensors, consider practical applications and do not aim for fundamental understanding, thus falling into 2) Pure applied research. Some of these application-oriented papers, about the creation and validation of nanodevices, also contribute to fundamental understanding placing them partly in category 3). This alignment of our scoring with Pasteur's theory validates our methodological contribution, which previous studies had not assessed.

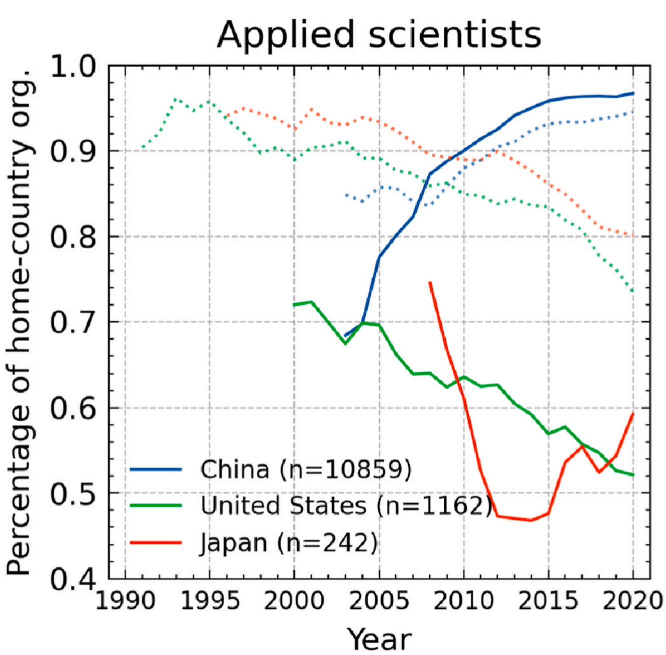

Fig. 6. Trends in the percentage of scientists affiliated with organizations in their home countries for the nanocarbon datasets. The trends are compared between applied scientists. The year of publication is set to 2010 or later because we want to focus on applied research that has characterized China since 2010. We calculated the percentage of authors affiliated or not affiliated with their home organization in the year they published their paper and plotted it as a 5-year moving average. Comparing China, the U.S., and Japan by nationality, China shows a different trend than the other two countries. The dotted lines are other groups of scientists randomly selected for comparison.

Table 2
Contribution of the main results.

Research question	Theoretical framework	Result and its value
How can basic and applied research be accurately quantified?	Pasteur's Quadrant	The identification results using the quantification indicator Topic-Aware Level Score (TALS), calculated based on large-scale data, can be interpreted through the theoretical framework. This contributes to the discussion on basic and applied research.
How does the research focus of a country shift from basic to applied research over time?	Linear Model	This study validates the advance from basic to applied research over 50 years using quantitative metrics, confirming global trends following the Linear Model.
How does scientific collaboration change during the transition from basic to applied science?	N/A	This study has enhanced understanding of international collaboration patterns by using quantitative metrics to analyze the shift from basic to applied research, demonstrating that local talents primarily fuel China's applied research sector.

Secondly, the transition of TALS over 50 years by country generally shows a shift from basic to applied research (Fig. 5b–d,f), quantitatively capturing the advance based on the Linear Model (Balconi et al., 2010; Bush, 2020). A common pattern in material science is that both the number of papers and the TALS are highest in China, indicating China's excellent momentum and advance in this field. Adams (2012) predicted in 2012 an increasing global trend in international co-authorships, and by 2020, anticipated a shift in the epicenter of science from the West to other regions, including China and India, which have indeed become increasingly influential. The shifts in country focus and the progression from basic to applied research provide an overview of field development. This perspective has become a valuable tool in funding, offering new criteria for decision-making and understanding one aspect of outcomes that was previously unavailable. For instance, Japan's decline in the number of papers on nanocarbon may be due to its inability to persistently invest in basic research (Fig. 5a). This study not only aligns with these meta-trends through publication counts but also through metrics that gauge the shift from basic to applied research, underscoring the substantial contributions of this research in reflecting global scientific dynamics.

Thirdly, this study successfully demonstrated patterns of international collaboration, which are influenced by social and geopolitical factors (Luukkonen et al., 1992). The proportion of Japanese and American researchers affiliated with domestic organizations is decreasing, indicating an increase in their involvement with foreign entities (Fig. 6). Conversely, in China, there is a growing trend of scientists affiliating more with domestic institutions. Researchers in Japan and the U.S. tend to engage more in basic research with lower TALS, while Chinese scientists are more involved in applied research (Fig. 5). This suggests that international collaboration is more prevalent in basic research, whereas applied research tends to have less, which aligns with the previous findings (Coccia and Wang, 2016).

Further, it turns out that the rise in applied research was not primarily due to returning scientists but rather supported by researchers who were already based domestically, although China's "Sea Turtle" policy aimed to attract researchers back home to boost its scientific output (Jonkers and Tijssen, 2008). This agrees with the opinion of experts who are at the center of nanocarbon research. Though it's widely recognized that China is expanding its influence in global science, when it comes to applied research, the scientists driving its growth are predominantly based within the country. The concentration might be a result of many non-academic entities like companies that have been growing by taking advantage of nanotechnology flourishment in China (Xin and Yidong, 2006; Zhou and Leydesdorff, 2006; Kostoff, 2012; Dong et al., 2016). By utilizing quantitative metrics for basic to applied research, this study has achieved a more nuanced understanding of these international collaboration patterns, revealing how local talents predominantly drive the applied research sector in China.

As a methodological implication, we successfully developed a novel framework to be applicable to any field other than biomedical. Traditional methods for scoring the progression of academic research from basic to applied stages relied on pre-defined and categorized terminology datasets linked to the target papers, resulting in restrictions on the use of biomedical fields only (Ke, 2019). We removed the constraint by employing topic models to automatically extract categorized terms from the large-scale paper dataset. Applying our framework to nanocarbon material papers, the calculated score aligned well with journal information, expert insights, and meta-trends (Figs. 3–5). Therefore, we expect that the method could generally measure the scientific advance from basic to applied research. By embedding the automatically extracted terms into a vector space, theoretically, we can spatially represent the meaning that the terms hold. The expression of a gradient from basic to applied research in latent space can enable us to obtain the capability to evaluate and visualize the potential of numerous domains.

Our methodological framework provides an objective way to assess how close the research is to practical application for decision-makers in funding agencies like government officials and corporate executives who allocate budgets for science and technology. The advantages of our framework are objectivity and comprehensiveness. For objectivity, conducting expert interviews across all the disciplines to be invested is not cost-effective. The methodology proposed here can act as a substitute for expert interviews when gauging the maturity level of applied research in various fields. These indicators capture comprehensive trends. While conventional metrics like publication count and citation numbers can highlight the level of attention a field is receiving, they do not indicate how close it is to application. By visualizing the trends in our metric, one can gain a more nuanced understanding of these trajectories. It is expected that this data-driven metric can provide an antenna-like function. The approach contributes to data utilization in funding (Lane and Bertuzzi, 2011), which will become increasingly important during transition periods of STI policy (OECD, 2023).

This method has several limitations. The definitions of basic and applied research can vary depending on the individual, era, and field. Despite simplifying human annotation, when creating a dictionary using topic modeling, there might be reliance on such subjectivity. A possible solution could be integrating patent citation information to hold objectivity. This should be considered in future work.

Fields such as advanced materials, biotechnology, and electronic devices, which require extensive basic research to ensure the reproducibility of scientific results, are well-suited to capture the flow from basic to applied research using our framework. On the other hand, in some fields, only basic or applied research may be conducted without any connection between the two. For example, artificial intelligence is a field that exemplifies this. Since deep learning has emerged, it has had a tremendous impact on society. However, the underlying principles, including the mathematically rigorous theory behind its high performance, are not yet fully understood. Applied research has advanced significantly, but basic research that seeks fundamental understandings of why this is possible still awaits further progress. In such fields, it may not be suitable to apply our framework to capture advances from basic to applied research.

Other fields might exhibit a flow from applied back to basic research (McKelvey, 1985). However, the flexibility of this method enables extracting paper data that spans from applied to basic research. By labeling each topic, we might define an axis from applied to basic in the semantic space, resulting in quantitative understanding. For example, it may visualize a flow that the improvement of scanning electron microscopes leads to atomic-level material surface science.

For scalability, our approach opens the door for meta-level insights into the dynamics of science. By using Topic-Aware Level Score to categorize researchers and their papers into either basic research or applied research, we can offer novel perspectives that may not only inform investment decisions but also contribute to emerging academic areas like Science of Science. The novel perspective, Topic-Aware Level Score may give us further opportunities to understand societal structures in academia. It could be worth considering blending with recently developed indicators like novelty and disruptive index (Liu et al., 2023). Regarding the uniqueness of the research environment in China, a thorough investigation into the movement of scientists and the appearance of co-authorship networks would provide a better understanding.

CRediT authorship contribution statement

Noriyuki Higashide: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yi Zhang:** Supervision, Validation, Writing – review & editing. **Kimitaka Asatani:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Takahiro Miura:** Data curation, Writing – review & editing. **Ichiro Sakata:** Funding acquisition, Supervision, Writing – review & editing.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve readability and language by generating paraphrase candidates for sentences. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data availability

Some data will be available on request but some is not because of no permission.

Acknowledgements

The author thank all members who have discussed on this work for invaluable comments. This work was supported by JSPS KAKENHI Grant Number 21K19817.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.technovation.2024.103050.

References

Adams, J., 2012. The rise of research networks. Nature 490, 335–336.
Ahmadpoor, M., Jones, B.F., 2017. The dual frontier: patented inventions and prior scientific advance. Science 357, 583–587.
Albert, M., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Res. Pol. 20, 251–259.
Álvarez Bornstein, B., Bordons, M., 2021. Is funding related to higher research impact? exploring its relationship and the mediating role of collaboration in several disciplines. Journal of Informetrics 15, 101102.
Asatani, K., Mori, J., Ochi, M., Sakata, I., 2018. Detecting trends in academic research from a citation network using network representation learning. PLoS One 13, 1–13.
Asatani, K., Oki, S., Momma, T., Sakata, I., 2023. Quantifying progress in research topics across nations. Sci. Rep. 13, 4759.
Athey, S., 2017. Beyond prediction: using big data for policy problems. Science 355, 483–485.
Baas, J., Schotten, M., Plume, A., Côté, G., Karimi, R., 2020. Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. Quantitative Science Studies 1, 377–386.
Balconi, M., Brusoni, S., Orsenigo, L., 2010. In defence of the linear model: an essay. Res. Pol. 39, 1–13.
Begum, P., Ikhtiari, R., Fugetsu, B., 2011. Graphene phytotoxicity in the seedling stage of cabbage, tomato, red spinach, and lettuce. Carbon 49, 3907–3919.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.
Boyack, K.W., Klavans, R., Börner, K., 2005. Mapping the backbone of science. Scientometrics 64, 351–374.
Boyack, K.W., Patek, M., Ungar, L.H., Yoon, P., Klavans, R., 2014. Classification of individual articles from all of science by research level. Journal of Informetrics 8, 1–12.
Bozeman, B., Corley, E., 2004. Scientists' collaboration strategies: implications for scientific and technical human capital. Res. Pol. 33, 599–616.
Bozeman, B., Youtie, J., 2016. The Strength in Numbers: How to Create Science Collaboration Dream Teams. Princeton University Press, Princeton, NJ.
Bozeman, B., Fay, D., Slade, C.P., 2013. Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. J. Technol. Tran. 38, 1–67.
Bozeman, B., Gaughan, M., Youtie, J., Slade, C.P., Rimes, H., 2016. Research collaboration experiences, good and bad: dispatches from the front lines. Sci. Publ. Pol. 43, 226–244.
Bush, V., 2020. Science, the Endless Frontier. Princeton University Press.
Cao, Q., Li, Y., Peng, H., 2023. From university basic research to firm innovation: diffusion mechanism and boundary conditions under a u-shaped relationship. Technovation 123, 102718.
Chu, H., Hu, X., Wang, Z., Mu, J., Li, N., Zhou, X., Fang, S., Haines, C.S., Park, J.W., Qin, S., Yuan, N., Xu, J., Tawfick, S., Kim, H., Conlin, P., Cho, M., Cho, K., Oh, J., Nielsen, S., Alberto, K.A., Razal, J.M., Foroughi, J., Spinks, G.M., Kim, S.J., Ding, J., Leng, J., Baughman, R.H., 2021. Unipolar stroke, electroosmotic pump carbon nanotube yarn muscles. Science 371, 494–498.
Coccia, M., 2009. What is the optimal rate of r&d investment to maximize productivity growth? Technol. Forecast. Soc. Change 76, 433–446.
Coccia, M., 2017. Sources of technological innovation: radical and incremental innovation problem-driven to support competitive advantage of firms. Technol. Anal. Strat. Manag. 29, 1048–1061.
Coccia, M., 2018a. General properties of the evolution of research fields: a scientometric study of human microbiome, evolutionary robotics and astrobiology. Scientometrics 117, 1265–1283.
Coccia, M., 2018b. Optimization in r&d intensity and tax on corporate profits for supporting labor productivity of nations. J. Technol. Tran. 43, 792–814.
Coccia, M., 2020. The evolution of scientific disciplines in applied sciences: dynamics and empirical properties of experimental physics. Scientometrics 124, 451–487.
Coccia, M., 2023. New perspectives in innovation failure analysis: a taxonomy of general errors and strategic management for reducing risks. Technol. Soc. 75, 102384.
Coccia, M., Bozeman, B., 2016. Allometric models to measure and analyze the evolution of international research collaboration. Scientometrics 108, 1065–1084.
Coccia, M., Roshani, S., 2023. General laws of funding for scientific citations: how citations change in funded and unfunded research between basic and applied sciences. J. Data Inform. Sci. 0.
Coccia, M., Roshani, S., 2024. Research funding and citations in papers of nobel laureates in physics, chemistry and medicine, 2019-2020. J. Data Inform. Sci. 8, 1–23.
Coccia, M., Wang, L., 2016. Evolution and Convergence of the Patterns of International Scientific Collaboration, vol. 113. Proceedings of the National Academy of Sciences, pp. 2057–2061.
Coccia, M., Watts, J., 2020. A theory of the evolution of technology: technological parasitism and the implications for innovation magement. J. Eng. Technol. Manag. 55, 101552.
Dong, H., Gao, Y., Sinko, P.J., Wu, Z., Xu, J., Jia, L., 2016. The nanotechnology race between China and the United States. Nano Today 11, 7–12.
Du, J., Li, P., Guo, Q., Tang, X., 2019. Measuring the knowledge translation and convergence in pharmaceutical innovation by funding-science-technology-innovation linkages analysis. Journal of Informetrics 13, 132–148.
Dudley, J.M., 2013. Defending basic research. Nat. Photonics 7, 338–339.
Fleming, L., Greene, H., Li, G., Marx, M., Yao, D., 2019. Government-funded research increasingly fuels innovation. Science 364, 1139–1141.
Goji, T., Hayashi, Y., Sakata, I., 2020. Evaluating "startup readiness" for researchers: case studies of research-based startups with biopharmaceutical research topics. Heliyon 6, e04160.
Grover, A., Leskovec, J., 2016. node2vec: scalable feature learning for networks. KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 855-864.
Guimera, R., Uzzi, B., Spiro, J., Amaral, L.A.N., 2005. Team assembly mechanisms determine collaboration network structure and team performance. Science 308, 697–702.
He, X., Htoon, H., Doorn, S.K., Pernice, W.H.P., Pyatkov, F., Krupke, R., Jeantet, A., Chassagneux, Y., Voisin, C., 2018. Carbon nanotubes as emerging quantum-light sources. Nat. Mater. 17, 663–670.
Higashide, N., Yoshida, M., Uda, T., Ishii, A., Kato, Y.K., 2017. Cold exciton electroluminescence from air-suspended carbon nanotube split-gate devices. Appl. Phys. Lett. 110, 191101.
Higashide, N., Asatani, K., Sakata, I., 2023. Domestically interacted scientists drive nanocarbon applied research in China. In: 2023 Portland International Conference on Management of Engineering and Technology (PICMET), pp. 1–7.
Hutchins, B.I., Davis, M.T., Meseroll, R.A., Santangelo, G.M., 2019. Predicting translational progress in biomedical research. PLoS Biol. 17, 1–25.
Ibhadon, A.O., Fitzpatrick, P., 2013. Heterogeneous photocatalysis: recent advances and applications. Catalysts 3, 189–218.
Islam, N., Miyazaki, K., 2009. Nanotechnology innovation system: understanding hidden dynamics of nanoscience fusion trajectories. Technol. Forecast. Soc. Change 76, 128–140. Knowledge Driven Planning Tools for Emerging and Converging Technologies.
Islam, N., Miyazaki, K., 2010. An empirical analysis of nanotechnology research domains. Technovation 30, 229–237.
Jefferson, O.A., Jaffe, A., Ashton, D., Warren, B., Koellhofer, D., Dulleck, U., Ballagh, A., Moe, J., DiCuccio, M., Ward, K., et al., 2018. Mapping the global influence of published research on industry and innovation. Nat. Biotechnol. 36, 31–39.
Jiao, Y., Zheng, Y., Jaroniec, M., Qiao, S.Z., 2015. Design of electrocatalysts for oxygen- and hydrogen-involving energy conversion reactions. Chem. Soc. Rev. 44, 2060–2086.
Jonkers, K., Tijssen, R., 2008. Chinese researchers returning home: impacts of international mobility on research collaboration and scientific productivity. Scientometrics 77, 309–333.
Justino, C.I., Gomes, A.R., Freitas, A.C., Duarte, A.C., Rocha-Santos, T.A., 2017. Graphene based sensors and biosensors. TrAC, Trends Anal. Chem. 91, 53–66.
Ke, Q., 2019. Identifying translational science through embeddings of controlled vocabularies. J. Am. Med. Inf. Assoc. 26, 516–523.
Ke, Q., 2020a. An analysis of the evolution of science-technology linkage in biomedicine. Journal of Informetrics 14, 101074.
Ke, Q., 2020b. Technological impact of biomedical research: the role of basicness and novelty. Res. Pol. 49, 104071.
Kostoff, R.N., 2012. China/usa nanotechnology research output comparison—2011 update. Technol. Forecast. Soc. Change 79, 986–990.
Kyebambe, M.N., Cheng, G., Huang, Y., He, C., Zhang, Z., 2017. Forecasting emerging technologies: a supervised learning approach through patent analysis. Technol. Forecast. Soc. Change 125, 236–244.
Lane, J., 2010. Let's make science metrics more scientific. Nature 464, 488–489.
Lane, J., Bertuzzi, S., 2011. Measuring the results of science investments. Science 331, 678–680.
Lewison, G., Paraje, G., 2004. The classification of biomedical journals by research level. Scientometrics 60, 145–157.
Li, Q., Mahmood, N., Zhu, J., Hou, Y., Sun, S., 2014. Graphene and its composites with nanoparticles for electrochemical energy applications. Nano Today 9, 668–683.
Li, D., Azoulay, P., Sampat, B.N., 2017. The applied value of public investments in biomedical research. Science 356, 78–81.
Li, X., Xie, Q., Jiang, J., Zhou, Y., Huang, L., 2019. Identifying and monitoring the development trends of emerging technologies using patent analysis and twitter data mining: the case of perovskite solar cell technology. Technol. Forecast. Soc. Change 146, 687–705.
Liu, L., Jones, B.F., Uzzi, B., Wang, D., 2023. Data, measurement and empirical methods in the science of science. Nat. Hum. Behav. 7, 1046–1058.
Luukkonen, T., Persson, O., Sivertsen, G., 1992. Understanding patterns of international scientific collaboration. Sci. Technol. Hum. Val. 17, 101–126.
Macilwain, C., 2010. Science economics: what science is really worth. Nature 465, 682–684.
Manjunath, A., Li, H., Song, S., Zhang, Z., Liu, S., Kahrobai, N., Gowda, A., Seffens, A., Zou, J., Kumar, I., 2021. Comprehensive analysis of 2.4 million patent-to-research citations maps the biomedical innovation and translation landscape. Nat. Biotechnol. 39, 678–683.
Mansfield, E., 1991. Academic research and industrial innovation. Res. Pol. 20, 1–12.
Maruyama, S., 2002. A molecular dynamics simulation of heat conduction in finite length swnts. Phys. B Condens. Matter 323, 193–195. Proceedings of the Tsukuba Symposium on Carbon Nanotube in Commemoration of the 10th Anniversary of its Discovery.

Maruyama, S., Kojima, R., Miyauchi, Y., Chiashi, S., Kohno, M., 2002. Low-temperature synthesis of high-purity single-walled carbon nanotubes from alcohol. Chem. Phys. Lett. 360, 229–234.
McKelvey, J.P., 1985. Science and technology-the driven and the driver. Technol. Rev. 88, 38.
Michel, J., Bettels, B., 2001. Patent citation analysis.a closer look at the basic input data from patent search reports. Scientometrics 51, 185–201.
Miura, T., Asatani, K., Sakata, I., 2021. Large-scale analysis of delayed recognition using sleeping beauty and the prince. Applied Network Science 6, 48.
Mosleh, M., Roshani, S., Coccia, M., 2022. Scientific laws of research funding to support citations and diffusion of knowledge in life science. Scientometrics 127, 1931–1951.
Nakamura, H., Kajikawa, Y., Suzuki, S., 2013. Multi-level perspectives with technology readiness measures for aviation innovation. Sustain. Sci. 8, 87–101.
Narin, F., Breitzman, A., 1995. Inventive productivity. Res. Pol. 24, 507–519.
Narin, F., Pinski, G., Gee, H.H., 1976. Structure of the biomedical literature. J. Am. Soc. Inf. Sci. 27, 25–45.
Narin, F., Hamilton, K.S., Olivastro, D., 1997. The increasing linkage between u.s. technology and public science. Res. Pol. 26, 317–330.
Newman, M.E., 2001. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA 98, 404–409.
Newman, M.E., 2004. Coauthorship networks and patterns of scientific collaboration. Proc. Natl. Acad. Sci. USA 101, 5200–5205.
OECD, 2023. OECD Science, Technology and Innovation Outlook 2023.
Ogawa, T., Kajikawa, Y., 2015. Assessing the industrial opportunity of academic research with patent relatedness: a case study on polymer electrolyte fuel cells. Technol. Forecast. Soc. Change 90, 469–475.
Ogawa, T., Takeuchi, M., Kajikawa, Y., 2018. Analysis of trends and emerging technologies in water electrolysis research based on a computational method: a comparison with fuel cell research. Sustainability 10.
Pan, R.K., Kaski, K., Fortunato, S., 2012. World citation and collaboration networks: uncovering the role of geography in science. Sci. Rep. 2, 902.
Pielke Jr., R., 2012. Basic research as a political symbol. Minerva 50, 339–361.
Radicchi, F., Fortunato, S., Castellano, C., 2008. Universality of citation distributions: toward an objective measure of scientific impact. Proc. Natl. Acad. Sci. USA 105, 17268–17272.
Roshani, S., Bagherylooieh, M.R., Mosleh, M., Coccia, M., 2021. What is the relationship between research funding and citation-based performance? a comparative analysis between critical disciplines. Scientometrics 126, 7859–7874.
Scarrà, D., Piccaluga, A., 2022. The impact of technology transfer and knowledge spillover from big science: a literature review. Technovation 116, 102165.
Shibata, N., Kajikawa, Y., Sakata, I., 2010. Extracting the commercialization gap between science and technology — case study of a solar cell. Technol. Forecast. Soc. Change 77, 1147–1155.
Shibata, N., Kajikawa, Y., Sakata, I., 2011. Detecting potential technological fronts by comparing scientific papers and patents. Foresight 13, 51–60.
Sinatra, R., Wang, D., Deville, P., Song, C., Barabási, A.L., 2016. Quantifying the evolution of individual scientific impact. Science 354, aaf5239.
Soliman, N.A., Chanut, N., Deman, V., Lallas, Z., Ulm, F.J., 2020. Electric energy dissipation and electric tortuosity in electron conductive cement-based materials. Phys. Rev. Mater. 4, 125401.
Stokes, D.E., 2011. Pasteur's Quadrant: Basic Science and Technological Innovation. Brookings Institution Press.
Sugimoto, C.R., Robinson-Garcia, N., Murray, D.S., Yegros-Yegros, A., Costas, R., Larivière, V., 2017. Scientists have most impact when they're free to move. Nature 550, 29–31.
Sun, X., Kaur, J., Milojević, S., Flammini, A., Menczer, F., 2013. Social dynamics of science. Sci. Rep. 3, 1069.
Surkis, A., Hogle, J.A., DiazGranados, D., Hunt, J.D., Mazmanian, P.E., Connors, E., Westaby, K., Whipple, E.C., Adamus, T., Mueller, M., Aphinyanaphongs, Y., 2016. Classifying publications from the clinical and translational science award program along the translational research spectrum: a machine learning approach. J. Transl. Med. 14, 235.
Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q., 2015. Line: large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, pp. 1067–1077.
Uddin, S., Hossain, L., Rasmussen, K., 2013. Network effects on scientific collaborations. PLoS One 8, e57546.
Weber, G.M., 2013. Identifying translational science within the triangle of biomedicine. J. Transl. Med. 11, 126.
Weinberg, B.A., Owen-Smith, J., Rosen, R.F., Schwarz, L., Allen, B.M., Weiss, R.E., Lane, J., 2014. Science funding and short-term economic activity. Science 344, 41–43.
Xin, H., Yidong, G., 2006. China bets big on big science. Science 311, 1548–1549.
Yan, E., Wu, C., Song, M., 2018. The funding factor: a cross-disciplinary examination of the association between research funding and citation impact. Scientometrics 115, 369–384.
Yang, X., Feng, L., Yuan, J., 2023. Research on linkage of science and technology in the library and information science field. Data and Information Management 7, 100033.
Zhang, Y., Porter, A.L., Hu, Z., Guo, Y., Newman, N.C., 2014. "term clumping" for technical intelligence: a case study on dye-sensitized solar cells. Technol. Forecast. Soc. Change 85, 26–39.
Zhang, Y., Zhang, G., Chen, H., Porter, A.L., Zhu, D., Lu, J., 2016. Topic analysis and forecasting for science, technology and innovation: methodology with a case study focusing on big data research. Technol. Forecast. Soc. Change 105, 179–191.
Zhang, Y., Lu, J., Liu, F., Liu, Q., Porter, A., Chen, H., Zhang, G., 2018. Does deep learning help topic extraction? a kernel k-means clustering method with word embedding. Journal of Informetrics 12, 1099–1117.
Zhou, P., Leydesdorff, L., 2006. The emergence of China as a leading nation in science. Res. Pol. 35, 83–104.

NORIYUKI HIGASHIDE received the B.A. degree in electrical engineering from Chiba University, in 2013 and the M.A. degree in electrical engineering from The University of Tokyo, in 2016. With several years of experience as a data scientist at IBM and a start-up, he is currently a Ph.D. candidate in Technology Management for Innovation, Graduate school of Engineering, The University of Tokyo. His research in Technology Management, Science of Science, and Network Science seeks to bridge scientific knowledge with industry challenges. n.higashide@ipr-ctr.t.u-tokyo.ac.jp

YI ZHANG serves as a Senior Lecturer at the UTS Australian Artificial Intelligence Institute. He holds dual PhD degrees, one is in Management Science & Engineering from Beijing Institute of Technology, in 2016 and the other is in Software Engineering in from the University of Technology Sydney. 2017, and he was a visiting scholar at the School of Public Policy, Georgia Institute of Technology from 2011 to 2012. He has been recognized as the Australian Research Field Leader in Library & Information Science in The Australian's Research Awards 2023. Yi.Zhang@uts.edu.au

KIMITAKA ASATANI received the B.A. degree in material engineering and the M.A. degree in computer science, and the Ph.D. degree in computer science from The University of Tokyo, in 2015. He is currently an Specially Appointed Lecturer at The University of Tokyo. He studies complex networks and their applications, focusing particularly on social and knowledge relationships in scientific articles and social networking services. asatani@tmi.t.u-tokyo.ac.jp

TAKAHIRO MIURA received the B.A. and the M.A., Ph.D. degree in engineering from Technology Management for Innovation, Graduate school of Engineering, The University of Tokyo. His research interests include Scientometrics, Science of Science, Management of Technology, and Network Science. He is seeking to reveal the unknowns of the development of science through a data-driven approach to help our decision-making, especially focusing on delayed recognition in science, known as Sleeping Beauty or Slow-cited papers. miura@ipr-ctr.t.u-tokyo.ac.jp

ICHIRO SAKATA received the B.A. degree in economics from The University of Tokyo, Tokyo, Japan, in 1989, the M.A. degree in international economics and finance from Brandeis University, Boston, MA, USA, in 1997, and the Ph.D. degree in environmental and ocean engineering from The University of Tokyo, in 2003. He became a Full Professor, in 2008, after working with the Ministry of Economy, Trade and Industry. He has held several appointments at The University of Tokyo, including the positions of a special advisor to the President, from 2015 to 2019 the Director of the Policy Alternatives Research Institute, from 2014 to 2016 the Head of the Vision Formation Group, FSI, since 2016, and the Head of the Department of TMI, Graduate School of Engineering, from 2016 to 2017. He is currently a Special Advisor to the President and a Professor with the School of Engineering, The University of Tokyo. His research interests include innovation management, technological forecasting, strategic research planning, and computational social science. He proposed the concept of technology informatics. He is a member of the Engineering Academy of Japan. Regarding social contributions, he served as a Special Advisor for the Ministry. isakata@ipr-ctr.t.u-tokyo.ac.jp