Scientometrics
https://doi.org/10.1007/s11192-025-05245-x

Comparing examiner citations and applicant citations: insights into technology evolution

Yali Qiao[1] · Alan L. Porter[2] · Ying Huang[3,4] · Haiyun Xu[5] · Xuefeng Wang[6]

Received: 20 March 2018 / Accepted: 17 January 2025

Abstract
Patent citation data is widely used in the study of technology evolution, but existing research has overlooked an issue that there may be potential differences between examiner citations and applicant citations, which may introduce biases from examiner citations. Yet, there is still a lack of systematic comparative study on the differences between applicant citations and examiner citations for technology evolution. To address this, we conducted a comprehensive comparison using USPTO patent data across four dimensions: technology profiling, technology relevance, technology diversity, and technology evolution pathways. For our case study, we selected the promising research area of photovoltaic cells. After comparing nine sub-technologies in this area, we have drawn some conclusions: (1) Applicants tend to provide more citations than examiners, and examiners tend to cite more recent patents than applicants; (2) There is no apparent inclination for applicants to avoid citing particularly relevant patents. On average, examiner citations are slightly closer in technological proximity to their invention than those cited by applicants; (3) The degree of diversity for applicant citations, examiner citations, and applicant & examiner citations at a single patent level lacks consistency. However, their average trend by year or by sub-technology is similar after adding examiner citations; (4) Merging family members strongly impacts main pathways through added examiner citations, which is quite contrary in the citation network with only USPTO-granted patents without merging patent members; (5) In sub-technologies at the growth stage, applicants and examiners both cite more recent patents and tend to integrate border technologies from other fields, which can be used as an indicator for evaluating the potential to become emerging. The findings remind us to pay extra attention to the context in which citation data is used to measure technology evolution, and can serve as signals for technology assessment as well.

Keywords Patent citation analysis · Examiner citation · Applicant citation · Technology diversity · Main path analysis · Photovoltaic cells (PV)

Extended author information available on the last page of the article

Published online: 10 February 2025

Introduction

Over the past few decades, patent citation analysis has increasingly become a mainstay for measuring technology evolution because of its advantages in quantitatively describing the dynamic of knowledge flows and technological advances (Jaffe & Trajtenberg, 2002; Ko et al., 2014). From the pure analysis of citing and cited patents to co-citations and bibliographic couplings, patent citations have become a powerful indicator for various objectives (Boyack & Klavans, 2010). However, there is much controversy regarding the use of patent citation data. A crucial element to consider is that, unlike academic paper citations, the citations in a patent result from a highly mediated process involving the applicant, the patent attorney, and the patent examiner (Meyer, 2000). As a general rule, applicants need to declare their inventions to be novel. Yet, very few laws compel applicants to provide literature or consult their peers to prove the validity of these declarations (Garfield, 1966). Due to a legal requirement to justify the patentability of a patent (Criscuolo & Verspagen, 2008), examiners must fully retrieve prior art and find the difference for the claimed invention. This process sometimes results in citations being added to the patents (Tan & Roberts, 2010). The distinct sources and motives of patent citation data have raised several issues regarding whether to use the full citation dataset, especially concerning whether examiner citations introduce noise into the analysis.

By analyzing the distribution of examiner and applicant citations in terms of self-citations, distance, technology overlaps, and vintage, Alcacer and Gittelman (2006) suggest that inferences about applicant knowledge derived from pooled citations may be subject to biased statistical outcomes or inflated standard errors. Jaffe et al. (1993) argued that examiner citations should be considered "disturbing noise" in the measurement of knowledge spillovers, as they typically do not reflect the technical spillovers between applicants. In a study of European Patent Office patents, Criscuolo and Verspagen (2008) also suggested that applicant citations are better viewed as indicators of knowledge flow, rather than relying on the total set of citations. These serve as a reminder that it is crucial to distinguish between examiner citations and applicant citations.

Previous research comparing examiner citations and applicant citations has yielded some interesting findings regarding their significant differences involving geographical and institutional knowledge spillover, value and quality of citations, technical composition, proportion of each type, knowledge proximity, etc. By analyzing prior art listed in US patents granted in 2001–2003, Alcacer et al. (2009) find that examiners provide an average of 63% of the citations to a patent application, and around 40% do not even include a single citation from an applicant. Other studies, however, show that 41% of the citations originate from the examiners (Sampat, 2010). In fact, one of the major research problems in the field of patent citations is to map the reaches of the sphere from which examiners draw their citations – from which regions and disciplines? Over what time periods? etc. Thompsori (2006) reports that applicants prefer to cite local patents, while examiners tend to cite patents from more distant places. Similarly, patent examiners may be more likely than applicants to cite related technologies in different industries. Using the likelihood of renewal as the measure of a patent's value, Hegde and Sampat (2009) find that examiner citations to a patent are stronger predictors than the applicant. Further, Li et al. (2014) analyzed the different motivations behind the citations provided by examiners and applicants. By dividing the applicant citations into general citations and self-citations, they find that examiner citations indicate linkages in science more accurately than an applicant's general citations and more comprehensively than an applicant's self-citations. Yasukawa and Kano (2014) explored the use of examiner citations from the perspective of the applicants' self-selection decisions with Japanese patent applications, finding that for an application, "YES" to each decision point during the patent examination may lead to a significantly larger number of examiner citations.

In contrast, other research reports no obvious tendency toward using a specific citation type. Applicant citations stem from the science underlying technology, while examiner citations smooth the distribution of citations across technology classes (Azagra-Caro et al., 2011). Research on patents issued by the Spanish Patent and Trademark Office (OEPM) implies that indicators should be selected on a context-driven basis rather than by the type of citation to determine the real-world implications of any differences between them (Azagra-Caro et al., 2011). Indeed, patent examiners rarely use applicant-submitted art during their assessment process. Rather, they rely almost exclusively on prior art in which they find themselves (Cotropia et al., 2013). Further, Park et al. (2017) compared the technological value of patents provided by applicants and examiners, and their results indicate that, in most sectors, the quality of patents cited by applicants is higher than that of those cited by examiners, but in all sectors, examiners tend to cite newer patents.

Despite the valuable contributions of previous research, a key question in these comparative studies remains unanswered: What are the differences between the applicant and examiner citations for describing technology evolution? Are changes for knowledge flow made if examiner citations are added to applicant citations in patent citation analysis? As a critical aspect of current research on patent citations, a systematic comparative study of the differences in these data is essential for advancing our understanding of characterizing technology evolution.

In this paper, we attempt to explore the differences between examiner citations and applicant citations for describing technology evolution using a suite of quantitative analysis approaches from both macro and micro perspectives. Given that regulations concerning applicant and examiner citations differ across countries, this study primarily focuses on USPTO patents, with photovoltaic (PV) cells as the case study. Four dimensions-technology profiling, technology relevance, technology diversity, and technology evolution pathways are explored for detailed comparison. The first dimension is the basic details of citation information, such as the citation distribution and the average citation lag of the patents cited. The second and third dimensions are to view the technology integration and knowledge transfer of patent to patent from the micro-perspective of technological evolution. The last dimension is to view the technological changes in a whole technical field from the macro-perspective of technological evolution. This systematic comparison for insights into technology evolution will broaden our understanding of these two different types of patent citations.

The reminder of this paper is organized as follows. The "Related work" Section reviews prior literature to illustrate the theory of technology relevance, technology diversity, and technological evolution pathways with patent citations. In the section "Methods and Data", the proposed research framework, the detailed process of measurement, and data collection are presented. In the section "Results", the comparison results and applications are analyzed. Finally, the "Conclusions and Discussions" section discusses this research's findings, contribution, limitations, and future work.

Related work

In general, patent citation indicates a technological relationship between the focal patent and its cited prior art. Essentially, patent citations originate from the examination process for applying for a patent, in which its relevant prior arts in the corresponding technological field are required to be fully searched and added as citations to legally clarify the patentability and non-obviousness and demarcate the reach of the intellectual property (IP) rights and claims conferred by the patent (EPO, 2023a; USPTO, 2023). Both applicants and examiners can provide these prior related arts. Their shares and motives vary and thus might deliver different performances. Focusing on the technological relevance between the focal patent and its cited prior arts, patent citation information is widely used for representing technological advancement and evolution (Huang et al., 2021). Generally, they can be divided into three categories, as shown in Fig. 1. From a macro-perspective of technological evolution, patent citation networks in a whole area vividly draw the complex genealogies of technological changes and *evolution pathways* over a long period. From the micro perspective, patent citations reflect the *technology diversity* of its integrated prior knowledge, as well as the *technology relevance* between the invention and each citation.

Technology relevance and patent citations

Studies have shown that the relation of a patent with its citations can reveal technological knowledge flow from the cited patent to the focal patent. According to patent law, patent citations should be the invention's relevant prior art with higher similarity, involving technology background, similar invention, knowledge base, improvements on previous work, etc. (EPO, 2023b; WIPO, 2016). Based on this, citations are assumed to have high technology relevance with its claimed inventions. Under this default knowledge relatedness between citing and cited patents, it enables masses of investigations on technological knowledge diffusion from their technological antecedents to descendants through forward and backward citations involving diverse entities like applicants, institutes, countries, etc. (Jiang et al., 2022a; Miguelez & Temgoua, 2020; Rosell & Agrawal, 2009). Also, more complex bibliometric indicators like co-citation, coupling, and the relatedness of each pair's forward citation network (Rodriguez et al., 2015) are proposed to evaluate the technology similarity of two patents. Among these, there is an assumption hidden that all citations are equal and can appropriately represent the knowledge flow from prior invention to the claim. However, besides knowledge flow generated by applicants, citations can also be provided by examiners and for other purposes (like ones added by the 3rd party). Whether all citations show similar performance is still not clear. Yet, there is a lack of study on the difference between the technology relevance of different sources of citations and focal patents.

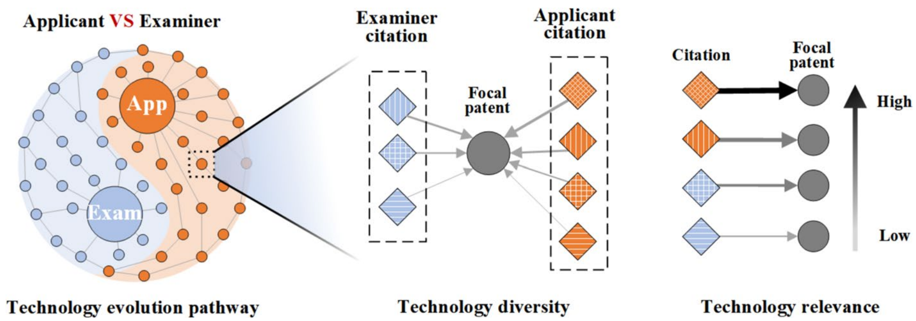


Fig. 1 The macro and micro perspective of technology evolution in view of patent citations

Technology diversity and patent citations

Like citations in research articles, for applicants, cited references for a patent reveal the technology learning and integration when developing a new invention (Ho et al., 2014). With the aid of bibliometrics, backward citation analysis is able to capture each invention's technological antecedents (Trajtenberg et al., 1997) and can clearly describe the flows and convergence mechanism from diverse knowledge sources (Kim & Lee, 2017). For its structured advantage, in citation analysis, technology diversity captures the knowledge breadth for a body of cited references as reflected by the range of subjects or patent classifications (Miller, 2006). Accordingly, the more dissimilar technical classes of citations for an invention involved, the higher integration degree of diverse knowledge sources. On the basis of this, researchers have widely used patent citations to assess technology diversity, ranging from assessing the integration ability of knowledge for research entities (Wang et al., 2022), to representing technology convergence with backward-citation information (Kim & Lee, 2017; Ko et al., 2014). Nonetheless, among these researches, the difference that examiner citations may bring is ignored.

Technology evolution pathways and patent citations

At a macro level, the forming of a patent citation network reflects the process of knowledge accumulation and technology advancement. By tracing the sources and targets for each node in the whole citation network, researchers can detect the process of technological evolution and predict future directions clearly and easily (Erdi et al., 2013). From the essence of technological evolution, technology interactions, and knowledge flows always follow certain patterns, but most patterns do not have the scope to push the development of a technology forward. Hence, to simplify and analyze the technological trajectories and identify core patents in a large and complex citation network, a mathematical tool-main path analysis (MPA) has frequently been introduced as the most effective tool to explore the structural backbone of a selected technical field (Huang et al., 2017). Currently, main path analysis is usually conducted on the whole citation network to represent the technology trajectory of concerned areas. Few researchers have uncovered the differences between applicant and examiner citations and the possible impact that examiner-added citations may generate.

Overall, from the perspective of technological evolution, patent citation analysis has become the most attractive method of detecting knowledge flows, technology diversity, and tracing technology pathways. However, a key issue has been ignored in most current studies when using patent citation analysis, which is quite different from citations in research articles: citations in patent documents can be provided by applicants, and can also be provided by examiners. For the purpose of patent examination, it is still unclear whether the citation data added by examiners will have an impact on the analysis of knowledge flows and technology evolution. There is also a lack of systematic comparative studies on applicant citations and examiner citations from the perspective of technological evolution. To this end, this study starts from the insights into technology evolution, comprehensively and systematically compares the differences between applicant and examiner citations with multiple dimensions, and further explores the possible impact by examiner-added citations.

Methods and data

Research framework and method

To explore the differences between patent examiner citations and applicant citations for characterizing technology evolution, we developed the research framework shown in Fig. 2. First, the USPTO-granted patent dataset with citation data is collected and divided into three sub-datasets according to their different citation sources: applicant citations, examiner citations, and applicant & examiner citations. Second, basic technology profiling analysis with the number of cited patents, the coverage of citations in focal patents, the average cited patent numbers per patent, and the average citation lag are calculated and compared. Third, the relevance of technology between focal patents and citations is measured by both the relevance category and the IPC4/CPC4-based similarity analysis for different citation sources. Fourth, True Diversity (TD) is used to estimate the technology diversity for each type of citation. Finally, main path analysis (MPA) is introduced to identify core patents and technological trajectories for comparison. The measurement details of technology relevance, technology diversity, and technology evolution pathways are illustrated below.

Technology relevance

In this research, the percentage of each citation category (which is tagged in the search report during the examination stage) and their cosine similarity-based IPC4/CPC4 are calculated to measure the technology relevance between cited patents and focal patents.

At the World Intellectual Property Organization (WIPO), it is recommended that documents referred should be tagged with letters “X” and “Y” to indicate particular relevance, and letters “A” “D” “E” “L” “O” “P” “T” to indicate other relevant prior art (WIPO, 2016).

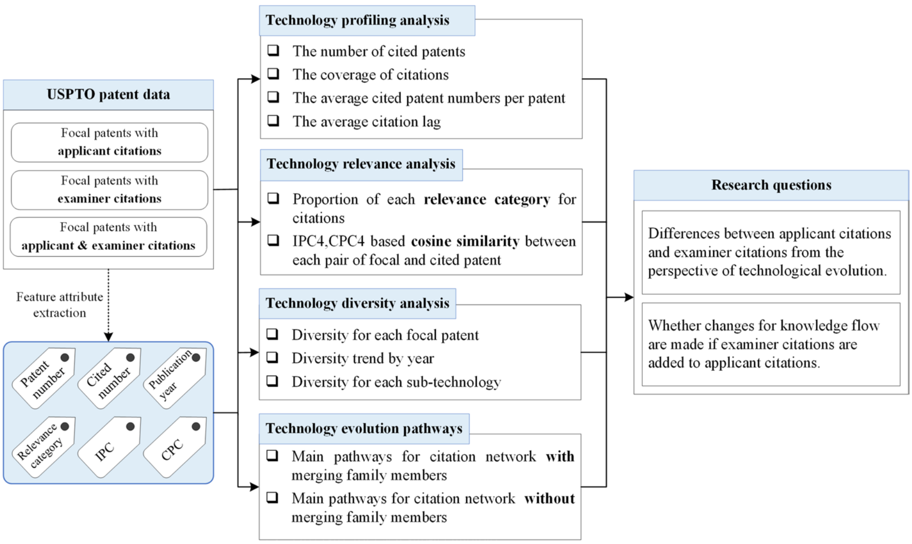


Fig. 2 The research framework

At EPO and other offices like JPO, patent documents cited in search reports are assigned specific or combinations of the relevant categories of X, Y, P, A, D, etc. (EPO, 2023b). Particularly, only in the examination stage of PCT patents for international protection at USPTO, cited documents will be categorized as X, Y, A, O, P, T, E, D according to their relevance by following guidelines by WIPO (USPTO, 2021). A detailed definition of each relevant category is depicted in Table 1. To compare the difference of citation relevance with focal patents for applicants and examiners, we calculate the percentage of each citation category for the whole technical area, that is, the proportion of cited patents marked with a certain category to all cited patents marked by applicants or examiners.

It is well known that the patents are allocated one or more IPCs and CPCs according to the different areas of technology to which they pertain. We further matched each reference to its corresponding IPC-4/CPC-4 information. According to Ejermo (2005), technology classes are widely considered an appropriate analysis unit for measuring technological proximity. We then calculated the similarity distributions between each patent and its cohort applicant-cited patents, examiner-cited patents, and applicant & examiner-cited patents in terms of cosine distance.

Technology diversity

Prior studies have used a variety of diversity indicators at the individual paper level. As Porter et al. (2008) declared, the citations contained in a single paper arguably reflect some degree of knowledge integration by the paper's author(s) of knowledge drawn from references cited. Originally, diversity represented in a paper's citations was evaluated using an integration score of the cognitive distance between the technological classification of the journals that were cited, such as the Web of Knowledge subject categories (Porter et al., 2007) or the Web of Science categories (WCs) (Porter & Rafols, 2009). In this paper for patent documents, we have used IPCs instead of WCs to gauge the degree of diversity in

Table 1 The definition of the relevance category

Relevance category	Definition
A	Technology background
D	Document cited in the application
E	Earlier patent document but published on, or after, the filing date
I	Particularly relevant if taken alone, prejudicing inventive step (EP, from April 2011)
L	Document cited for other reasons
P	EP—intermediate document; WO—Published prior to the international filing date but later than the priority date claimed
T	Theory or principle underlying the invention
X	EP—Particularly relevant if taken alone; WO- Document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
Y	EP—Particularly relevant if combined with another document of the same category; WO—Document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more such documents, such combination being obvious to a person skilled in the art

the technology fields cited. Rather than applying the relatively common Rao-Stirling diversity (Stirling, 2007), we used the measure called "True Diversity (TD)" because of its superior discriminatory power (Zhang et al. (2016). The formulation of True Diversity (TD) is:

$$TD = \frac{1}{\sum_{i \neq j} p_i p_j \left(1 - d_{ij}\right)}$$

Here, x_i is the number of references to the *i-th* category, and $p_i = {}^{x_i}/_{X}$;$X = \sum x_i$. d_{ij} is the dissimilarity between the categories i and j.$d_{ij} = 1 - s_{ij}$, where, s_{ij} is the cosine similarity between the categories i and j.

Technology evolution pathways

This research uses main path analysis (MPA) to trace the evolution of technology pathways. The main pathways are the paths from the source to the sink with the highest traversal weights (De Nooy et al., 2005). In the pathways, we only considered data from the most prominent graph components and disregarded patents that neither cited nor were cited by at least one other patent—in other words, "orphan" patents digressing from the mainstream of the domain. This reduced set of patents was configured into a merged citation network and then examined using main path analysis based on bibliographic and/or patent citations to identify the most significant and constructive patents over the history of each technology.

When conducting main path analysis, three factors should be taken into consideration: (1) how to construct the citation network for patents; (2) how to measure the traversal weight of each citation link from a set of starting vertices to the end vertex; and (3) how to select the highest traversal weights to map out the most representative path. There are two main ways to construct a citation network for patents: by relationships of individual patents and by relationships of patent families. Their difference is described in Fig. 3. The widely used DWPI patent family standards have grouped patents that directly share at least one priority and basically the same technical content into a family (Jiang et al., 2022b). Therefore, taking advantage of patent family information, a citation network with merging patent family members can better capture the technology evolution structure among different technical topics (Huang et al., 2021). For a comprehensive comparison, this paper explores the role of applicants and examiner citations by constructing citation networks using both two ways.

As for the traversal weight of each citation link, in this paper, we followed Liu et al. (2019)'s recommendations and applied search path link count (SPLC) as the preferred choice for traversal weight because it describes knowledge diffusion scenarios in science and technology development better than the other traversal weights. Traditionally, the chosen algorithms are "local, " meaning they start with a node and then consistently follow the link with the highest traversal count. However, the local main path approach only highlights the significance of local progress, but a global main path approach is developed instead to estimate the overall importance of each node in knowledge flow globally (Liu & Lu, 2012). Based on this, we used the global MPA method to obtain the paths that have been the most significant in terms of overall traversal count for all citation networks. It is worth noting that forward, backward, and standard in the global-search-based method may overlook some links with the highest traversal weight, and we use the global key-route MPA method introduced by Liu and Lu (2012) to identify the main pathways for each citation network (Liu et al., 2013). Following these steps above, main pathways of applicant

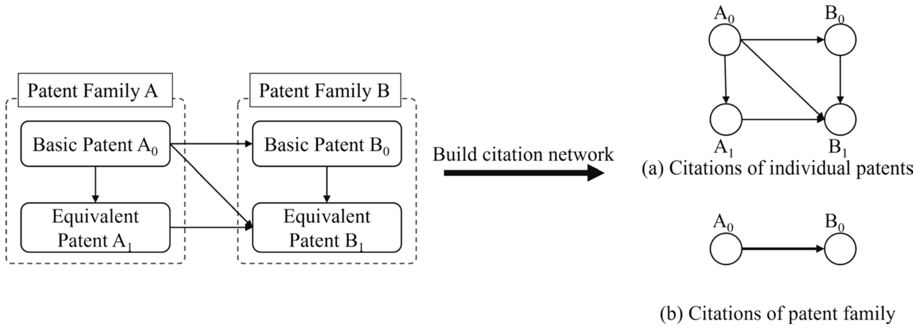


Fig. 3 The difference between the two ways to construct a citation network

citations, examiner citations and applicant & examiner citations are mapped to trace the role of applicants and examiners.

Data retrieval and collection

When the applicants and their attorneys submit an invention application to USPTO, they are obliged to provide a list of the references they consider to be state-of-the-art in terms of their relevance to the patentability of the invention. Further, USPTO patent citations added by examiners are clearly delineated with an asterisk on the front page of the issued patent, making these data a good choice for our analysis. We collected USPTO-granted patent data from Derwent Innovation because this platform brings together comprehensive international patent coverage and powerful intellectual property analysis tools.

Photovoltaics is a promising research frontier for future sustainability. It has specific advantages as an energy source and has received much attention over several decades (Shibata et al., 2010). However, searching for a domain with an abundance of common terminologies, such as 'Big Data' or 'Artificial Intelligence', can be especially challenging. Sometimes, finding relevant research on a very technical topic is easier than one that involves many broad and familiar concepts (Huang et al., 2015). Hence, the USPTO and EPO introduced a system of cooperative patent classifications (CPC), which indexes patents that focus on emerging technologies against specific tags in a new Y-class – for example, the tag Y02: Climate Change Mitigating Technologies (Leydesdorff et al., 2015). To find the photovoltaics patents for our sample, we searched for the term "CPC=(Y02E10/54*)" and, given that prior to 2001, applicant citation information was an optional recommendation (Alcacer et al., 2009), we searched for patent data between 1 January 2001 and 31 December 2021, setting the endpoint to one year prior to when we conducted the search to account for the time lag between the application and publication. Ultimately, 14,632 records were retrieved. The number of patents in each of the nine sub-technologies that make up photovoltaics is shown in Table 2.

Additionally, we charted the number of granted patents in photovoltaics and its nine sub-technologies broken down by year, as shown in Fig. 4. There was a steady increase in the number of patents from 2009 to a peak in 2015, followed by a remarkable decline in most sub-technologies. Differently, patents in Organic PV have been seeing continuous growth after 2015 due to the urgent need for renewable energy, especially for the promising future of solar energy. Also, most of the sub-technologies (except for *Organic PV*) show an obvious decrease

Table 2 The search results for nine sub-technologies of photovoltaics

CPC Code	Sub-technologies	Abbrev	Records
Y02E 10/541	CuInSe2 material PV cells	CuInSe2	1122
Y02E 10/542	Dye-sensitized solar cells	DSSCs	1222
Y02E 10/543	Solar cells from Group II-VI materials	Group II-VI	323
Y02E 10/544	Solar cells from Group III-V materials	Group III-V	1290
Y02E 10/545	Microcrystalline silicon PV cells	Microcrystalline	182
Y02E 10/546	Polycrystalline silicon PV cells	Polycrystalline	465
Y02E 10/547	Monocrystalline silicon PV cells	Monocrystalline	2768
Y02E 10/548	Amorphous silicon PV cells	Amorphous	1026
Y02E 10/549	Organic PV cells	Organic PV	7999

Some patents are tagged in more than a single sub-technology

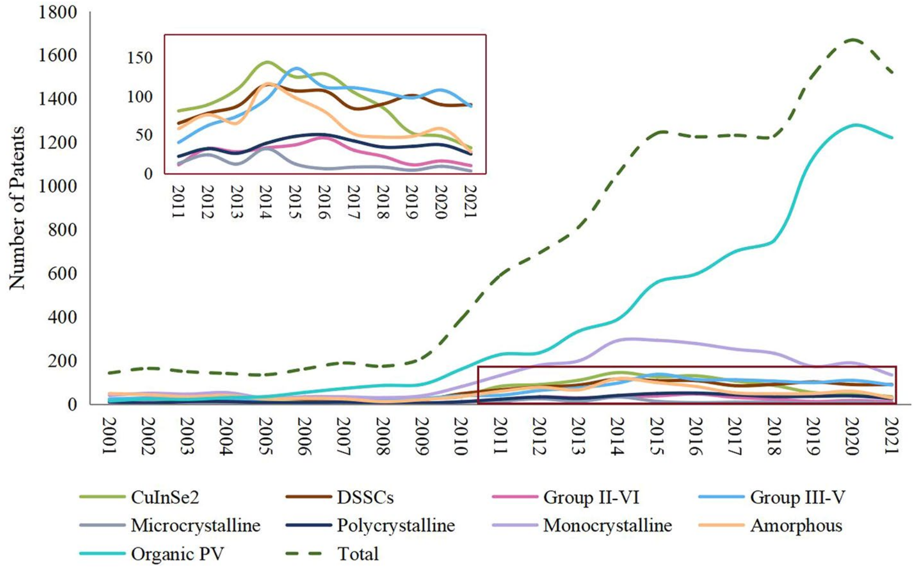


Fig. 4 The number of patents granted in the nine sub-technologies of photovoltaics

in patent activity. This is an explicit signal that photovoltaic technology has entered its maturity stage and may continue to wane unless there are some significant breakthroughs in the field. Unlike new & emerging science and technologies (NESTs) that are radically novel, relatively fast-growing, and characterized by a certain degree of coherence (Rotolo et al., 2017), mature technologies have a long and sufficiently rich history of citations to make exploring citation patterns relatively easy.

Results

We conducted four separate analyses—technology profiling analysis, technology relevance analysis, technology diversity analysis, and technology evolution pathway analysis to compare the citations from patent examiners and applicants. Each is discussed in turn in the next subsections.

Technology profiling analysis

Any citation in a patent could have been added by one of three parties: the applicant(s), the examiner(s), or a third party. Focusing on the citations added by applicants and examiners in the photovoltaics area, a total of 123,538 patent citations were added by applicants across 14,234 focal patents, while examiners added 54,881 patent citations across 14,287 focal patents. According to the patent law of EPO and USPTO, examiners impact patent citations significantly because they can add, retain, or even delete the original applicant's citations (EPO, 2023b; USPTO, 2023). For this, we explored these patterns in more detail by examining the distribution of each citation type over nine sub-technologies of photovoltaics, as shown in Table 3. From these results, we have found that most patents in photovoltaics contain citations that were added by either the applicant or the examiner, but the applicants added far more citations than the examiners in all sub-technologies (through NCP and ACP analysis), indicating that applicants contributed more citations to the total citation set. Yet, in most sub-technologies, the examiner citations had slightly wider coverage than the applicants' (through CCFP analysis).

The final step of the analysis involved extracting the publication year of the cited patents and focal patents. The average citation lag for each group was then calculated, providing insights into how far back applicants and examiners search prior art. The results are presented in Fig. 5. Overall, applicants tend to cite older patents, indicating that they search for prior art further back in time. In contrast, examiners tend to cite more recent patents (refer to Fig. 5(a)). The patterns remain consistent across all nine sub-technologies, as shown in Fig. 5(b). Specifically, in sub-technologies in the growth stage, such as Organic PV, both applicants and examiners cite more recent patents compared to other sub-technologies in the maturity stage.

Table 3 The citation distribution pattern of nine sub-technologies of photovoltaics

Sub-technology	Applicant			Examiner		
	NCP	CCFP	ACP	NCP	CCFP	ACP
CuInSe2	38,216	87.39%	31.71	7725	93.28%	6.41
DSSCs	18,112	86.15%	13.56	5954	84.21%	4.46
Group II-VI	13,641	89.62%	37.27	2010	90.44%	5.49
Group III-V	53,582	85.28%	37.55	8158	91.24%	5.72
Microcrystalline	6346	89.86%	29.24	1243	93.55%	5.73
Polycrystalline	12,685	88.04%	23.71	3119	91.40%	5.83
Monocrystalline	100,985	88.59%	32.64	17,548	92.05%	5.67
Amorphous	27,886	88.22%	24.7	6026	93.89%	5.34
Organic PV	193,891	90.79%	22.9	43,952	88.44%	5.19

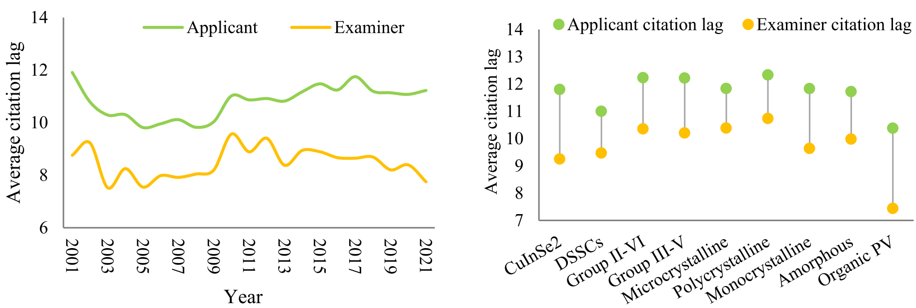
NCP: The number of cited patents; *CCFP:* The coverage of citations in focal patents; *ACP:* The average cited patent numbers per patent

Technology relevance analysis

Derwent Innovation has categorized citations based on their relevance for patent documents and search reports, and it could be used to mark the purpose and importance of citations. We summarized the citations added by the applicants and examiners in each category. Considering that except for PCT patents, USPTO-authorized patenting does not require labeling the relevant category of citations, we have also counted the distributions of each category for their family members.

As shown in Table 4, two important types of citations stand out: documents of particular relevance that restrict the claims of the applicant (citation categories X and Y); and references related to the technology's background (citation category A). Category A, which provides the prior art used in the steps toward an invention up to the current state-of-the-art, has a higher percentage of both applicant citations and examiner citations. Surprisingly, applicants contribute more citations to Categories X and Y. This is contrary to the intuition that applicants might deliberately leave out certain citations in an attempt to get broader patents. A higher proportion of X citations among the sample of applicant citations to examiner citation (26.00% vs. 25.82% or 8.21%) indicates that applicants have no obvious inclination to avoid citing patents that "already show essential features of the invention". Similarly, the Y category, which also includes patents compromising novelty but only in combination with other patents, occurs almost as frequently in the applicant citations.

To compare the relevance of citations with focal patents, we calculated the cosine similarity based on IPC4 and CPC4 between citations and focal patents for each applicant and examiner citation. The average results are presented in Table 5. Upon initial observation, it appears that patent examiners tend to cite sources more closely related to their invention than those cited by applicants, particularly in photovoltaics. However, there is significant variation in the degree of similarity in nine sub-technologies. Additionally, a comparison of applicant citations and citations from both applicants and examiners indicates that the average similarity scores increase when examiner citations are included.



(a) The average citation lag by year

(b) The average citation lag by sub-technology

Fig. 5 The average citation lag in the field of photovoltaics by year and by sub-technology

Table 4 The corresponding citation count of the relevance category of citations

Relevance Category	USPTO grant		Family members	
	No. of Cited Patents: Applicant	No. of Cited Patents: Examiner	No. of Cited Patents: Applicant	No. of Cited Patents: Examiner
A	19 (4.57%)	6784 (60.78%)	4521 (49.77%)	36,717 (51.9%)
D	-	-	311 (3.42%)	1080 (1.53%)
E	-	-	84 (0.92%)	733 (1.04%)
I	-	-	566 (6.23%)	3167 (4.48%)
L	-	-	0 (0.00%)	62 (0.09%)
P	-	-	357 (3.93%)	1444 (2.04%)
T	-	-	7 (0.08%)	43 (0.06%)
X	130 (31.25%)	916 (8.21%)	2362 (26.00%)	18,266 (25.82%)
Y	267 (64.18%)	3469 (31.08%)	2737 (30.13%)	16,425 (23.22%)

The percentage of each category by applicants or examiners is calculated by the proportion of cited patents marked with the certain category to all cited patents marked by applicants or examiners

Table 5 The average similarity between focal patents and cited patents

Sub-technology	IPC4 distributions			CPC4 distributions		
	Applicant	Examiner	Applicant & Examiner	Applicant	Examiner	Applicant & Examiner
CuInSe2	0.529	0.570	0.544	0.437	0.425	0.427
DSSCs	0.776	0.795	0.792	0.708	0.763	0.725
Group II-VI	0.788	0.801	0.803	0.743	0.761	0.745
Group III-V	0.504	0.520	0.520	0.449	0.475	0.453
Microcrystalline	0.668	0.667	0.678	0.637	0.683	0.654
Polycrystalline	0.786	0.822	0.811	0.703	0.773	0.729
Monocrystalline	0.770	0.775	0.784	0.715	0.756	0.728
Amorphous	0.728	0.769	0.749	0.697	0.748	0.708
Organic PV	0.797	0.797	0.811	0.716	0.770	0.736
Total	0.609	0.643	0.627	0.537	0.557	0.541

Applicant & Examiner means the cited patents are provided by applicants or examiners

Technology diversity analysis

From this insight into the diversity of technology fields cited, we attempted to further explore the differences in diversity between the two types of citations, as well as the impact of applicant citations on diversity after adding examiner citations. Figure 6 shows the technological diversity at IPC4 of the patents cited by each focal patent. The horizontal axis represents the degree of diversity for applicant citations, while the left and right vertical axes represent the degree of diversity for examiner citations and applicant & examiner citations, respectively, which the color of the nodes can distinguish. Nine sub-technologies have all shown that there is no consistency between the degree of diversity for applicant citations and examiner citations at the level of individual patents. Furthermore, the diversity for applicant & examiner citations also shows that when examiner citations are added, some patent citations increase in diversity while others decrease. This apparent difference between diversity for applicant citations, and applicant & examiner citations illustrates the large impact of examiner citations. Thus, we can conclude that examiner citations could impact the measurement of patent citation diversity in terms of knowledge flows for individual patents.

Figure 7 illustrates the variation in diversity between examiner and applicant citations over time. Overall, diversity values for both types of citations increased continuously from 2001 to 2008, followed by a decrease to a relatively stable level after 2009. Although no

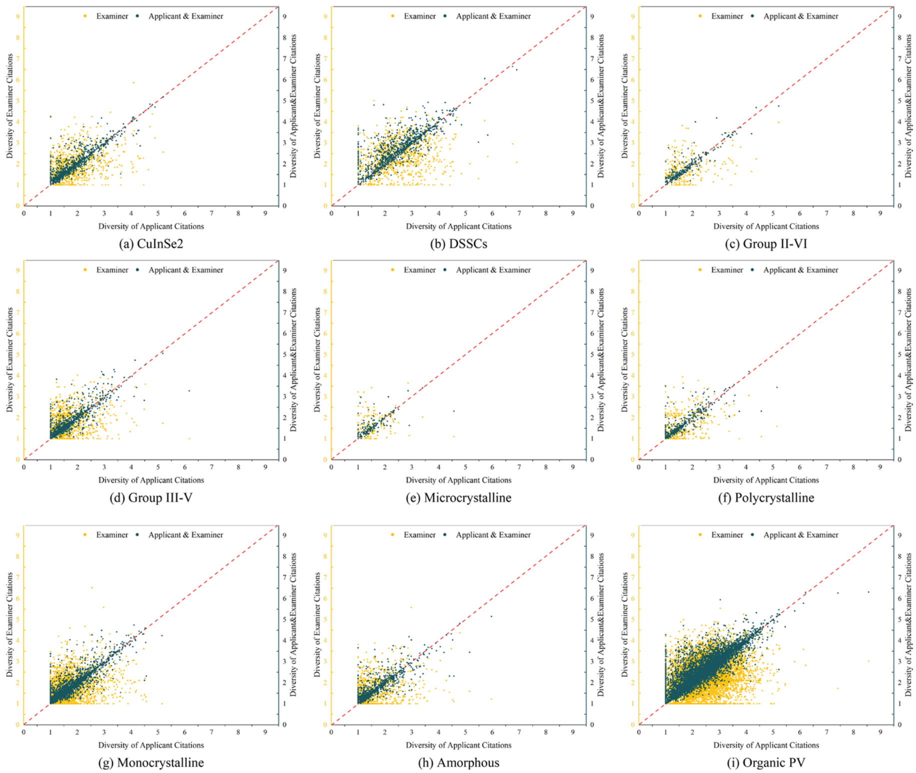


Fig. 6 The IPC4-based technology diversity of different citation sources at individual focal patent level

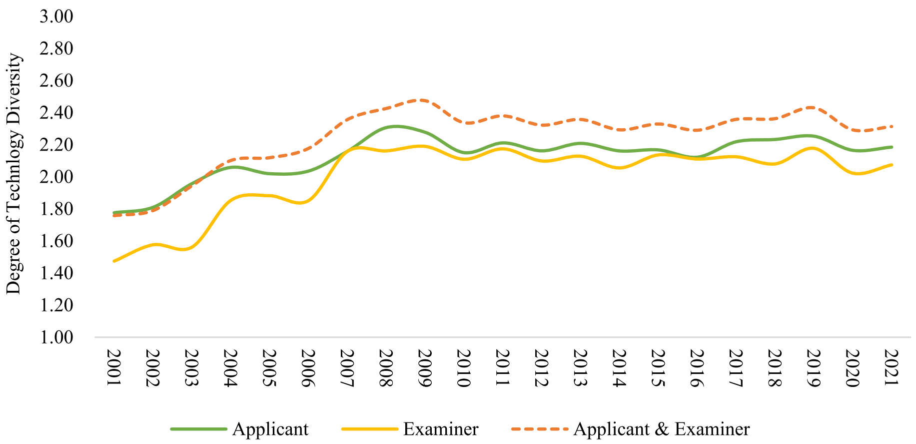


Fig. 7 The technology diversity at IPC-4 of the patents cited by year

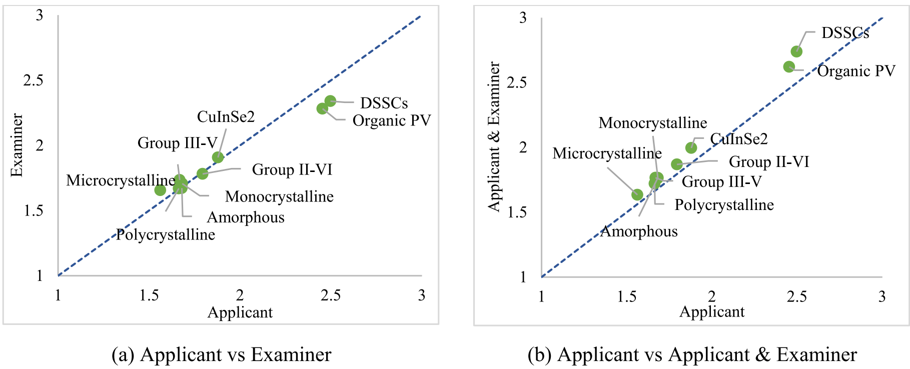


Fig. 8 The average degree of technology diversity in cited patents by sub-technology

obvious cyclic patterns are present, these trends indicate that in the early stages of photovoltaics, incorporating knowledge from outside the field was more crucial than in later stages of development. Additionally, the diversity value for applicant citations is consistently higher than that of examiner citations throughout the period. Although applicants contribute more citations to the total pool, the reasons for searching for prior art differ between applicants and examiners. Therefore, the combined scores show a higher diversity value than the scores for each alone. Additionally, the trend for applicant citations and the trend for combined applicant and examiner citations are quite similar. After examining the diversity of citations in the technical field for the year, it appears to have little impact on the measured trends of technology diversity when examiner citations are added.

Figure 8 illustrates the average degree of diversity per patent by sub-technology. By comparing the diversity of applicant citations and examiner citations, it is found that most sub-technologies have higher diversity scores for applicant citations than examiner citations. Although examiner citations generally do not track with applicant citations, they tend to select related patents as references with similar characteristics. With the exception of *Organic PV* and *Dye-sensitized Solar Cells (DSSCs)*, which integrate border technologies into their fields, the other subclasses seem to cite patents in closer technologies. Patent citation data has a goldmine of new insights into the development of technologies, as the links between citation patterns can be viewed as technology aggregations and spillovers (Erdi et al., 2013). To some extent, the stronger the ability of *Organic PV* and *Dye-sensitized Solar Cells (DSSCs)* to integrate with other sub-technologies, the more potential they have to become emerging photovoltaic technologies. Moreover, results have also shown a similar trend of average diversity degree for each sub-technology with adding examiner citations; thus, this impact can be ignored when comparing different sub-technologies.

Main path analysis

Patent citation analysis approaches are good at structuring many patents to profile a technological or industrial landscape or capture knowledge transfers and changes within those landscapes (Huang et al., 2021). The network of patents connected through citations provides a representation of the innovation process (Erdi et al., 2013). Technology transfer patterns and knowledge flow can generally be divided into *interior citation* and *exterior citation*. This paper primarily focuses on interior citations, which demonstrate the relationships among patents within a specific technology rather than patents in other fields. After removing isolated nodes, we utilized social network analysis to compare the structures of three patent citation networks: citations by the applicant, citations by the examiner, and citations by both the applicant and examiner. Table 6 lists the comparative results of the network structures. The network constructed by applicant citations exhibits better connectivity with more nodes and arcs, and a shorter path length compared to the network constructed by examiner citations.

As depicted in the "Methods and Data" section, merging patent family members may make a difference in the generation of main pathways. Therefore, citation networks with and without merging family members are built for the following comparison. On the basis of the extracted largest connected sub-network for each citation network, main pathways for three types of citation networks are constructed respectively to trace the role of applicant and examiner citations in the photovoltaics field.

Ultimately, our analysis revealed three main pathways without merging family members by setting the key-route parameter as default value 1–10, consisting of 18 patents from

Table 6 The indicators of network analysis of three kinds of photovoltaic patent citation networks

Network structure indicators	Applicant citations	Examiner citations	Applicant & Examiner citations
Number of nodes	10,144	9970	12,349
Number of arcs	40,926	17,228	56,955
Average degree	4.035	1.728	4.579
Average path length	4.896	4.509	5.998
Modularity	0.765	0.793	0.742
Number of communities	303	448	179
Number of weakly connected components	286	422	165
Number of strongly connected components	9774	9847	11,819
Average clustering coefficient	0.080	0.027	0.082

the applicant & examiner citation network, 35 patents for the applicant citation network, and 24 patents for the examiner citation network. These networks, shown in Fig. 9, were generated by the large network analysis program *Pajek* (http://mrvar.fdv.uni-lj.si/pajek) and mapped in the science of science software *ITGInsight* (http://en.itginsight.com).

We can see large differences between the applicant citation network and the examiner citation network, with only 4 nodes (US6198091, US6198092, US6278055, US6352777) appearing in both main paths. Nonetheless, by comparing the main pathways of applicant citations, and applicant & examiner citations, it is found that there is little impact after adding examiner citations to applicant citations, with only 1 node (US6844025) of main pathways in the examiner citation network and 1 new node (US6844025) appearing, the other 16 nodes are all nodes in main pathways of an applicant citation network. This strongly indicates that applicants have a strong incentive to disclose known prior art, which plays a vital role in knowledge flow and technological learning in various ways. As for the 2 nodes that do not appear in the main pathways of the applicant citation network, further exploration is needed to judge whether they provide valuable clues for tracing the evolution of technology.

The patents on the main pathway may be the latent and significant cornerstones of a technology and, as such, may possibly play an important role in a technology's innovation. Hence, Table 7 provides some statistical information about each node in the main pathways of the applicant & examiner citation networks. In the early stages of a technology,

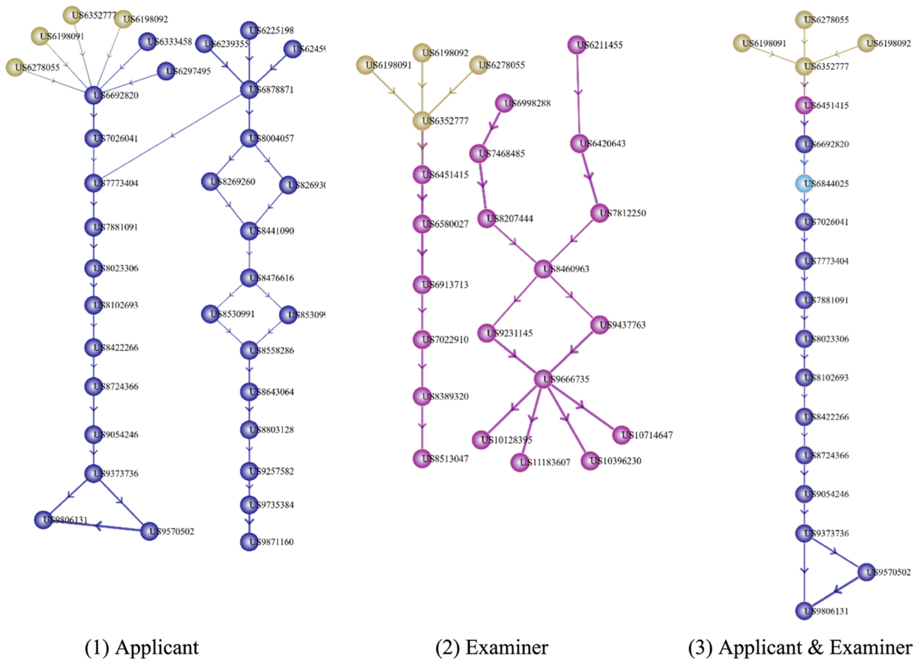


Fig. 9 The main pathways in the three types of photovoltaic citation networks (without merging family members). The brown nodes represent patents belonging to both main paths of the applicant citation network and examiner citation network; The dark blue nodes and pink nodes indicate these patents are only included in the main paths of the applicant citation network and examiner citation network, respectively; The light blue nodes represent that only appear in the applicant & examiner citation network

patent applications typically receive relatively more citations than the inventions that follow, even though most include less cited patents, which both applicants and examiners usually cite. For the nodes after US7026041 in the table, they cite a lot of patents, but most are added by the applicants, which is the main reason why the cited patents added by the examiner no longer appear in the main path of the applicant & examiner network. In addition, the patents that integrate a large number of related patents or diffuse knowledge to many others tend to increase the degree-centrality and show further potential to be tracked as important nodes in a technology development path. In terms of the most recent nodes, e.g., US8724366, US9054246, US9373736, US9570502, and US9806131, it is hard to tell whether they can be treated as critical nodes until more patents enter the field. Either way, they stand out as indispensable technological nodes for revealing the current research frontier. Regarding US6844025 (appearing only in the applicant & examiner citation network), and US6451415 (appearing in the examiner citation network), we have noticed that they both receive higher forward citations, indicating a higher technological value as well.

Our analysis also revealed three main pathways with merging family members by setting key-route parameter as default value 1–10, consisting of 35 patents from the applicant & examiner citation network, 32 patents for the applicant citation network, and 33 patents for the examiner citation network, as shown in Fig. 10. We can see that no patents appear in either the main paths of the applicant citation network or the examiner citation network. Six nodes appear in both the applicant & examiner citation network and the applicant

Table 7 Patents in the main pathways of the applicant & examiner citation networks (without merging family members)

Publication Number	Publication Year	A	E	Count of Citing Patents	Count of Cited Patents		
					Applicant & Examiner	Applicant	Examiner
US9806131	2017	√		0	290	278	12
US9570502	2017	√		1	286	274	12
US9373736	2016	√		4	285	273	12
US9054246	2015	√		4	282	270	12
US8724366	2014	√		5	279	267	12
US8422266	2013	√		9	267	255	12
US8102693	2012	√		17	256	244	12
US8023306	2011	√		17	254	244	10
US7881091	2011	√		23	239	229	10
US7773404	2010	√		40	236	224	12
US7026041	2006	√		38	45	44	1
US6844025	2005			88	16	15	1
US6692820	2004	√		148	15	14	1
US6451415	2002		√	311	12	7	5
US6352777	2002	√	√	352	21	14	7
US6278055	2001	√	√	221	16	12	4
US6198092	2001	√	√	225	16	15	1
US6198091	2001	√	√	239	16	13	3

A and E indicate that the patents are included in the main technology trajectory of the patent citation networks based on the applicant and examiner citations, respectively citation network, and six nodes appear in both the applicant & examiner citation network and the examiner citation network. Similarly, nodes that appear both in the main paths of the applicant & examiner citation network and examiner citation network are all patents in the early period, while nodes that appear both in the main paths of the applicant & examiner citation network and applicant citation network are more recent patents, because of the large number of cited patents by applicants.

The difference between the main paths of the applicant citation network and the applicant & examiner citation network, with merging family members, strongly indicates the great impact generated by added examiner citations. As for whether it is positive or negative, further comparison of nodes in main paths of the applicant & examiner citation network, and applicant citation network is needed to judge whether examiner citations provide valuable clues for tracing the evolution of a technology. Some statistical information about each node in the main pathways of the applicant & examiner citation networks and the applicant citation network after merging family members are presented in Table 8 and Table 9 in Appendix, respectively.

Through analysis, it is not difficult to find that patents with a higher proportion of examiner citations are more likely to appear both in the examiner citation network and the applicant & examiner citation network. Furthermore, compared with MPA in the applicant citation network, MPA nodes in the applicant & examiner citation network receive higher forward citations. When we include cited patent members belonging to a family from other patenting authorities, the greater impact is generated in the main pathways of the applicant & examiner citation network, indicating that examiners tend

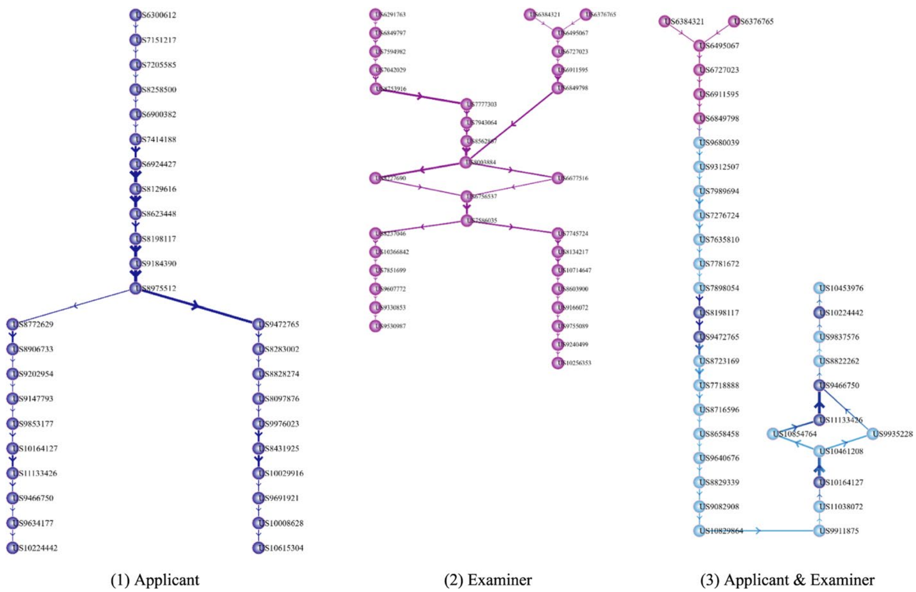


Fig. 10 The main pathways in the three types of photovoltaic citation network (merging family members). The brown nodes represent patents that belong to both main paths of the applicant citation network and examiner citation network; The dark blue nodes and pink nodes indicate these patents are included in the main paths of the applicant citation network and examiner citation network, respectively; The light blue nodes represent that only appear in the applicant & examiner citation network

Table 8 Patents in the main pathways of the *applicant & examiner* citation networks (merging family members)

Publication Number	Publication Year	A	E	Count of Citing Patents	Count of Cited Patents		
					Applicant & Examiner	Applicant	Examiner
US6384321	2002		√	99	4	1	3
US6376765	2002		√	197	9	2	7
US6495067	2002		√	107	8	0	8
US6727023	2004		√	18	4	2	2
US6911595	2005		√	44	27	3	24
US6849798	2005		√	13	4	0	4
US9680039	2017			0	97	96	1
US9312507	2016			4	221	215	6
US7989694	2011			5	44	22	22
US7276724	2007			110	43	41	2
US7635810	2009			119	104	95	9
US7781672	2010			58	74	38	36
US7898054	2013			0	227	227	0
US8198117	2012	√		27	197	193	4
US9472765	2016	√		0	43	42	1
US8723169	2014			4	47	45	2
US7718888	2010			60	14	5	9
US8716596	2014			17	13	6	7
US8658458	2014			5	8	8	0
US9640676	2017			3	19	11	8
US8829339	2014			8	37	24	13
US9082908	2015			1	14	14	0
US10829864	2020			1	241	238	3
US9911875	2018			1	9	0	9
US11038072	2021			0	251	246	5
US10164127	2018	√		0	538	532	6
US10461208	2019			0	69	14	55
US10854764	2020			0	35	6	29
US11133426	2021	√		0	101	101	0
US9935228	2018			0	13	7	6
US9466750	2016	√		0	39	36	3
US8822262	2014			16	87	83	4
US9837576	2017			4	32	23	9
US10224442	2019	√		2	31	24	7
US10453976	2019			0	17	2	15

A and E indicate that the patents are included in the main technology trajectory of the patent citation networks based on the applicant and examiner citations, respectively

to cite more patents from other authorities than applicants, which is consistent with the finding by Thompsori (2006). A family of patents is a collection of patents from different countries that are concerned with the same technical topic (Michel & Bettels,

Table 9 Patents in the main pathways of the *applicant* citation networks (merging family members)

Publication Number	Publication Year	A	A&E	Count of Citing Patents	Count of Cited Patents		
					Applicant & Examiner	Applicant	Examiner
US6300612	2001	√		311	17	12	5
US7151217	2006	√		17	40	39	1
US7205585	2007	√		5	47	45	2
US8258500	2012	√		2	16	10	6
US6900382	2005	√		61	110	94	16
US7414188	2008	√		7	154	146	8
US6924427	2005	√		45	131	115	16
US8129616	2012	√		9	36	30	6
US8623448	2014	√		1	167	164	3
US8198117	2012	√	√	27	197	193	4
US9184390	2015	√		3	4	2	2
US8975512	2015	√		0	128	128	0
US8772629	2014	√		0	135	132	3
US8906733	2014	√		0	104	100	4
US9202954	2015	√		3	106	103	3
US9147793	2015	√		5	28	26	2
US9853177	2017	√		1	12	9	3
US10164127	2018	√	√	0	538	532	6
US11133426	2021	√	√	0	101	101	0
US9466750	2016	√	√	0	39	36	3
US9634177	2017	√		1	12	12	0
US10224442	2019	√	√	2	31	24	7
US9472765	2016	√	√	0	43	42	1
US8283002	2012	√		2	34	33	1
US8828274	2014	√		0	28	21	7
US8097876	2012	√		2	33	31	2
US9976023	2018	√		1	30	27	3
US8431925	2013	√		5	44	27	17
US10029916	2018	√		1	126	122	4
US9691921	2017	√		5	139	132	7
US10008628	2018	√		0	179	179	0
US10615304	2020	√		0	184	183	1

2001). However, given the different citation behaviors of parent and child patents, global technology trends cannot be understood by only analyzing patent data issued by a single authority. All the patents cited within a complete family need to be merged into one family record to produce a full and comprehensive citation analysis.

Conclusions and discussion

There has been a lengthy discussion regarding the distinctions between applicant and examiner citations and whether the latter may cause ’noise’ for knowledge flow and spillover. This paper primarily focuses on the area where citation analysis is frequently employed—technology evolution—to comprehensively investigate the differences between examiner citations and applicant citations. After analyzing USPTO patent data from both micro and macro perspectives, we have drawn four conclusions based on our four-dimensional analysis:

- First, the proportion of examiner citations is lower than applicant citations, mainly because USPTO is the only authority requiring that applicants submit all known prior art documents; otherwise, they would be considered as violating the "duty of candor." This point contradicts the results of many previous studies, which are mostly based on EPO patent data (Criscuolo & Verspagen, 2008). In view of citation lag, the average ages of cited patents added by examiners are younger than those provided by the applicants. This can be attributed to the nature of the examination process, in which examiners prioritize the most recent patents to assess the patentability and originality of a patent application.
- Second, as for technology relevance analysis, the distribution of relevance category reveals that, for applicants, there is no obvious inclination to avoid citing particularly relevant patents, which is in line with the proportion of non-similar citations (the similarity value is 0) among the sample of applicant citations to examiner citation. Inconsistently, IPC/CPC-based average cosine similarity reveals that examiner citations are slightly closer in technological proximity to their invention than those that applicants cite. This is due to the fact that, in the examination process, their focus is on ensuring that the invention is distinct from prior art and represents an original contribution to the field, and therefore they are more likely to cite patents directly related to the technological aspects of the invention.
- Third, technology diversity analysis has shown that there is no consistency between the degree of diversity for applicant citations, examiner citations, and applicant & examiner citations at the single patent level. The average trends by year and by sub-technology reveal that the citations provided by examiners tend to be limited to a relatively narrow set of technological fields. This finding supports the view that examiners have the potential to increase the quality of citations to prior art, given their real-world experience and substantial knowledge of the field they are examining. A great impact is generated for individual patents with added examiner citations for the diversity measurement. Yet, there is little impact on the average trend of technology diversity by year or by sub-technology after adding examiner citations.
- Fourth, results without merging family members have presented that examiner citations cannot reverse the main pathways of an applicant citation network. They do contribute some unique nodes that have the potential to extend the process of technological innovation in a technological field. Such findings fit with the observation that examiners usually have rich experience in searching and retrieval and are familiar with professional patent databases. However, this is quite contrary in the citation network after merging family members, with a strong impact generated by added examiner citations. This needs our attention and further exploration when analyzing the main pathways in the citation network.

- Particularly, comparison among different sub-technologies has shown that in sub-technologies at the growth stage, applicants and examiners both cite patents more recently and tend to integrate border technologies from other fields than in sub-technologies at the mature stage for their rapid turnover, which, in turn, can be used as an indicator for evaluating the potential to become emerging.

The main contribution of this research is the provision of a systematic comparison between the two types of citations within the context of technology evolution, along with new insights from both macro and micro perspectives. Practically, these differences yield two important implications for effectively utilizing patent citation data. First, these differences remind us to pay extra attention to the context in which citation data is used to measure technology evolution, particularly when deciding whether to distinguish between applicant citations and examiner citations, and which one is more accurate. We have noticed that, when characterizing technology relevance and technology diversity, there are significant differences between applicant and examiner citations at the individual patent level. However, when comparing across fields or exploring annual trends, the impact of examiner citations is minimal for their consistency. When characterizing the main pathways of technology evolution, examiner citations had a significant impact in merging family member groups. Therefore, differences between the two types should not be ignored when using citation analysis, and whether to use the examiner citations needs careful consideration according to the actual problem to be solved. Second, the differences in patterns between examiner and applicant citations can serve as signals for technology assessment. For example, applicants and examiners both cite patents more recently and tend to integrate border technologies from other fields in sub-technologies at the growth stage than at the mature stage. This pattern can, in turn, serve as an indicator for evaluating the potential to become emerging. The fact that examiner citations tend to refer to younger patents suggests that examiner-driven citations may be more indicative of cutting-edge or emerging technologies.

However, this paper also has some limitations, which serve as fodder for further exploration. This study presents some interesting findings, but the reasons behind certain dimensional differences remain unclear. For example, the reason for the significant differences caused by examiner citations in the main pathways after merging family members remains unclear, which requires further in-depth investigation in the future. Further, our analysis focused on USPTO-granted patent data. However, given the different citation regulations of different patenting authorities, and therefore the differences in citation patterns in EPO, JPO, and other popular offices need to be compared further. Second, we came to realize that pooled citations provide a good signal of knowledge flows, but citation noise can be added at many stages of the patent process: by the applicants when preparing the application, by researchers during the patent office search, by the examiner during the examination phase, or by interested third parties after the patent application has been published (Derek, 2014). However, most of these citations are not reported separately in some intellectual property departments/offices, e.g., the Chinese patents in CNIPA (the China National Intellectual Property Administration) don't cover such information. For this, deep and detailed exploration is still needed in future work. At the same time, not all US examiner citations are made purely by examiners, in that some "examiner" citations in USPTO may have been recognized and recorded through information added by the information disclosure statement supplied by applicants; at the same time, some applicant citations are added by patent attorneys rather than applicants (Wada, 2018). How to distinguish this detailed citation information remains to be further explored. Otherwise, nine sub-technologies in photovoltaics are selected as case studies for analysis, and their comparisons have revealed some interesting findings compared to previous studies rather than analyzing multiple fields. In the future, more case studies should be conducted to bolster generalizability and to support or overturn our conclusions.

Appendix

See (Table 9).

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 72374162, L2324105), China Postdoctoral Science Foundation (Grant No. 2023M741901), the Wuhan University Library, and the National Laboratory Center for Library and Information Science at Wuhan University. The findings and observations presented in this paper are those of the authors and do not necessarily reflect the views of the supporters.

Funding National Natural Science Foundation of China, 72374162, L2324105, Ying Huang, China Postdoctoral Science Foundation, 2023M741901, Yali Qiao

References

Alcacer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics, 88*(4), 774–779. https://doi.org/10.1162/rest.88.4.774

Alcacer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in US patents: An overview and analysis. *Research Policy, 38*(2), 415–427. https://doi.org/10.1016/j.respol.2008.12.001

Azagra-Caro, J. M., Mattsson, P., & Perruchas, F. (2011). Smoothing the lies: The distinctive effects of patent characteristics on examiner and applicant citations. *Journal of the American Society for Information Science and Technology, 62*(9), 1727–1740. https://doi.org/10.1002/asi.21574

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science & Technology, 61*(12), 2389–2404. https://doi.org/10.1002/asi.v61:12

Cotropia, C. A., Lemley, M. A., & Sampat, B. (2013). Do applicant patent citations matter? *Research Policy, 42*(4), 844–854. https://doi.org/10.1016/j.respol.2013.01.003

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy, 37*(10), 1892–1908. https://doi.org/10.1016/j.respol.2008.07.011

De Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge University Press.

Derek, S. (2014). Finding the signal in the noise of patent citations: How to focus on relevance for strategic advantage. *Technology Innovation Management Review*, 4(9), 36–44. http://timreview.ca/article/830.

Ejermo, O. (2005). Technological diversity and Jacobs' externality hypothesis revisited. *Growth and Change, 36*(2), 167–195.

EPO (2023a). Guidelines for Examination, Part B-Guidelines for Search. https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_ii_4.htm. Accessed March, 2 2023.

EPO (2023b). Guidelines for Examination. https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_x_9_2.htm. Accessed March, 2 2023.

Erdi, P., Makovi, K., Somogyvari, Z., Strandburg, K., Tobochnik, J., Volf, P., et al. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics, 95*(1), 225–242. https://doi.org/10.1007/s11192-012-0796-4

Garfield, E. (1966). Patent citation indexing and the notions of novelty, similarity, and relevance. *Journal of Chemical Documentation, 6*(2), 63–65. https://doi.org/10.1021/c160021a001

Hegde, D., & Sampat, B. (2009). Examiner citations, applicant citations, and the private value of patents. *Economics Letters, 105*(3), 287–289. https://doi.org/10.1016/j.econlet.2009.08.019

Ho, M. H. C., Lin, V. H., & Liu, J. S. (2014). Exploring knowledge diffusion among nations: A study of core technologies in fuel cells. *Scientometrics, 100*(1), 149–171. https://doi.org/10.1007/s11192-014-1265-z

Huang, Y., Schuehle, J., Porter, A. L., & Youtie, J. (2015). A systematic method to create search strategies for emerging technologies based on the Web of Science: Illustrated for "Big Data." *Scientometrics, 105*(3), 2005–2022. https://doi.org/10.1007/s11192-015-1638-y

Huang, Y., Zhu, D. H., Qian, Y., Zhang, Y., Porter, A. L., Liu, Y. Q., et al. (2017). A hybrid method to trace technology evolution pathways: A case study of 3D printing. *Scientometrics, 111*(1), 185–204. https://doi.org/10.1007/s11192-017-2271-8

Huang, Y., Zhu, F. J., Porter, A. L., Zhang, Y., Zhu, D. H., & Guo, Y. (2021). Exploring Technology Evolution Pathways to Facilitate Technology Management: From a Technology Life Cycle Perspective. *IEEE Transactions on Engineering Management, 68*(5), 1347–1359. https://doi.org/10.1109/tem.2020.2966171

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics, 108*(3), 577–598. https://doi.org/10.2307/2118401

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT Press.

Jiang, L. D., Chen, J. Y., Bao, Y. H., & Zou, F. (2022a). Exploring the patterns of international technology diffusion in AI from the perspective of patent citations. *Scientometrics, 127*(9), 5307–5323. https://doi.org/10.1007/s11192-021-04134-3

Jiang, L. D., Zou, F., Qiao, Y. L., & Huang, Y. (2022b). Patent analysis for generating the technology landscape and competition situation of renewable energy. *Journal of Cleaner Production, 378*, 21. https://doi.org/10.1016/j.jclepro.2022.134264

Kim, J., & Lee, S. (2017). Forecasting and identifying multi-technology convergence based on patent data: The case of IT and BT industries in 2020. *Scientometrics, 111*(1), 47–65. https://doi.org/10.1007/s11192-017-2275-4

Ko, N., Yoon, J., & Seo, W. (2014). Analyzing interdisciplinarity of technology fusion using knowledge flows of patents. *Expert Systems with Applications, 41*(4), 1955–1963. https://doi.org/10.1016/j.eswa.2013.08.091

Leydesdorff, L., Alkemade, F., Heimeriks, G., & Hoekstra, R. (2015). Patents as instruments for exploring innovation dynamics: Geographic and technological perspectives on "photovoltaic cells." *Scientometrics, 102*(1), 629–651. https://doi.org/10.1007/s11192-014-1447-8

Li, R., Chambers, T., Ding, Y., Zhang, G., & Meng, L. S. (2014). Patent citation analysis: Calculating science linkage based on citing motivation. *Journal of the Association for Information Science and Technology, 65*(5), 1007–1017. https://doi.org/10.1002/asi.23054

Liu, J. S., & Lu, L. Y. Y. (2012). An integrated approach for main path analysis: Development of the Hirsch Index as an example. *Journal of the American Society for Information Science and Technology, 63*(3), 528–542. https://doi.org/10.1002/asi.21692

Liu, J. S., Lu, L. Y. Y., & Ho, M.H.-C. (2019). A few notes on main path analysis. *Scientometrics, 119*(1), 379–391. https://doi.org/10.1007/s11192-019-03034-x

Liu, J. S., Lu, L. Y. Y., Lu, W. M., & Lin, B. J. Y. (2013). Data envelopment analysis 1978–2010: A citation-based literature survey. *Omega-International Journal of Management Science, 41*(1), 3–15. https://doi.org/10.1016/j.omega.2010.12.006

Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. *Scientometrics, 49*(1), 93–123. https://doi.org/10.1023/a:1005613325648

Michel, J., & Bettels, B. (2001). Patent citation analysis - A closer look at the basic input data from patent search reports. *Scientometrics, 51*(1), 185–201. https://doi.org/10.1023/a:1010577030871

Miguelez, E., & Temgoua, C. N. (2020). Inventor migration and knowledge flows: A two-way communication channel? *Research Policy, 49*(9), 13. https://doi.org/10.1016/j.respol.2019.103914

Miller, D. J. (2006). Technological diversity, related diversification, and firm performance. *Strategic Management Journal, 27*(7), 601–619. https://doi.org/10.1002/smj.533

Park, I., Jeong, Y., & Yoon, B. (2017). Analyzing the value of technology based on the differences of patent citations between applicants and examiners. *Scientometrics, 111*(2), 665–691. https://doi.org/10.1007/s11192-017-2323-0

Porter, A. L., Cohen, A. S., David Roessner, J., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics, 72*(1), 117–147. https://doi.org/10.1007/s11192-007-1700-5

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics, 81*(3), 719. https://doi.org/10.1007/s11192-008-2197-2

Porter, A. L., Roessner, J. D., & Heberger, A. E. (2008). How interdisciplinary is a given body of research? *Research Evaluation, 17*(4), 273–282. https://doi.org/10.3152/095820208x364553

Rodriguez, A., Kim, B., Turkoz, M., Lee, J. M., Coh, B. Y., & Jeong, M. K. (2015). New multi-stage similarity measure for calculation of pairwise patent similarity in a patent citation network. *Scientometrics, 103*(2), 565–581. https://doi.org/10.1007/s11192-015-1531-8

Rosell, C., & Agrawal, A. (2009). Have university knowledge flows narrowed?: Evidence from Patent Data. *Research Policy, 38*(1), 1–13. https://doi.org/10.1016/j.respol.2008.07.014

Rotolo, D., Rafols, I., Hopkins, M. M., & Leydesdorff, L. (2017). Strategic intelligence on emerging technologies: Scientometric overlay mapping. *Journal of the Association for Information Science and Technology, 68*(1), 214–233. https://doi.org/10.1002/asi.23631

Sampat, B. N. (2010). When do applicants search for prior art? *Journal of Law & Economics, 53*(2), 399–416. https://doi.org/10.1086/651959

Shibata, N., Kajikawa, Y., & Sakata, I. (2010). Extracting the commercialization gap between science and technology — Case study of a solar cell. *Technological Forecasting and Social Change, 77*(7), 1147–1155. https://doi.org/10.1016/j.techfore.2010.03.008

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface, 4*(15), 707–719. https://doi.org/10.1098/rsif.2007.0213

Tan, D., & Roberts, P. W. (2010). Categorical coherence, classification volatility and examiner-added citations. *Research Policy, 39*(1), 89–102. https://doi.org/10.1016/j.respol.2009.11.001

Thompsori, P. (2006). Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. *Review of Economics and Statistics, 88*(2), 383–388. https://doi.org/10.1162/rest.88.2.383

Trajtenberg, M., Henderson, R., & Jaffec, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology, 5*(1), 19–50.

USPTO (2021). Patent Cooperation Treaty (PCT). https://www.uspto.gov/patents/basics/international-protection/patent-cooperation-treaty. Accessed March,2 2023.

USPTO (2023). Manual of patent examining procedure latest revision february 2023 [R-07.2022]: 707 Examiner's letter or action [R-07.2022]. https://mpep.uspto.gov/RDMS/MPEP/current#/current/d0e72880.html. Accessed March, 2 2023.

Wada, T. (2018). The choice of examiner patent citations for refusals: Evidence from the trilateral offices. *Scientometrics, 117*(2), 825–843. https://doi.org/10.1007/s11192-018-2885-5

Wang, Y., Qiao, Y. L., Wang, X. F., & Wan, D. (2022). Identifying the roles of research entities in technological knowledge flow among patents assignees by using patent citations. *IEEE Transactions on Engineering Management, 69*(6), 2754–2768. https://doi.org/10.1109/tem.2019.2956739

WIPO (2016) 'Handbook on industrial property information and documentation standard ST.14'. pp. 10–11. Available at: https://www.wipo.int/export/sites/www/standards/en/pdf/03-14-01.pdf.

Yasukawa, S., & Kano, S. (2014). Validating the usefulness of examiners' forward citations from the viewpoint of applicants' self-selection during the patent application procedure. *Scientometrics, 99*(3), 895–909. https://doi.org/10.1007/s11192-013-1195-1

Zhang, L., Rousseau, R., & Glänzel, W. (2016). Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. *Journal of the Association for Information Science and Technology, 67*(5), 1257–1265. https://doi.org/10.1002/asi.23487

Authors and Affiliations

Yali Qiao[1] · Alan L. Porter[2] · Ying Huang[3,4] · Haiyun Xu[5] · Xuefeng Wang[6]

✉ Ying Huang
ying.huang@whu.edu.cn

[1] Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China

[2] Search Technology, Inc., Norcross, GA 30092, USA

[3] Center for the Studies of Information Resources, School of Information Management, Wuhan University, Wuhan 430072, China

[4] Centre for R&D Monitoring (ECOOM) and Department of MSI, KU Leuven, 3000 Leuven, Belgium

[5] School of Information Management, Shandong University of Technology, Zibo 255000, China

[6] School of Management, Beijing Institute of Technology, Beijing 100081, China