Knowledge formation of MPEG: Analysis using bibliographic clustering of citation networks

TAMURA Suguru[1, 2]*, IWAMI Shino[3] and SAKATA Ichiro[2]

This study focused on the knowledge formation of an image-digitizing technology (MPEG). Based on the observed result, we propose the scenario for the knowledge creation including standardization. We firstly demonstrate how standardization changes the science linkage between patents and academic articles. Secondly, we discuss the relevance of the bibliographic clustering method (BCCN) for this analysis. Finally, the differences in knowledge creation related to standardization among countries are discussed.

Keywords : Science linkage, standardization, bibliographic clustering of citation networks

1 Introduction

In this study, we present the scenario to find more promising technology areas. We will use wide knowledge sources (i.e. academic articles, patents and standards). For the purpose, we will examine the knowledge formation of an image-digitizing technology (MPEG[Note 1)]). MPEG is one of the most influential standardized technologies in today's digital society because it is widely used for encoding and decoding audio and moving images.

We consider patents, academic papers, and standards as the factors for the discussion. In the analysis, these three factors prove to be the important elements to find promising technology areas. Moreover, the differences in knowledge creation related to standardization among countries such as China, Japan, and Korea, as well as Germany and the USA, are discussed.

Historically, MPEG has been standardized and developed under section JTC1/SC29 by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). We can, therefore, select the technology to study the influence of standardization on knowledge creation. Technologies of this kind have been widely used with regard to image recognition for the application of artificial intelligence technologies.[2]–[5] Such technologies are now becoming a prime R&D target[6] because these will contribute to social and organizational transformation.[7][8]

We examine data related to the MPEG technology using bibliographic information and observe the science linkage of the data in the exploration and exploitation process.[Note 2)] Using data analysis with computation (i.e., bibliographic clustering of citation networks (BCCN)) for patents, academic papers, and standards related to MPEG, we, to the best of our knowledge, are the first to empirically examine the influence of standardization on science linkage under this method. The results show that, for a specific technology, such as artificial intelligence-related imaging technology (i.e., MPEG), standardization affects science linkage and can change the knowledge flow between academic papers and patents.

The bibliographic clustering method is now widely used to study the knowledge flow and to mine new knowledge.[9] Several related studies have been conducted owing to the rapid development of Information and Communications Technology (ICT) and bibliographic data infrastructure.[10][11] The main research subject of such promising study is the relationship between patents and academic papers. Scholars call this knowledge flow "science linkage." Academic papers are important results of scientific research and are the source from which new technologies can be developed and applied to various fields.[12][13]

According to recent studies, standardization has both positive and negative influences on the production of patents and

---

1. Research Institute of Economy, Trade and Industry (RIETI)　1-3-1 Kasumigaseki, Chiyoda-ku 100-8901, Japan, 2. Institute for Future Initiatives (IFI), University of Tokyo　7-3-1 Hongo, Bunkyo-ku, 113-0033 Japan,　*E-mail: tamura.edu@nifty.com, 3. Institute of Laser Engineering, Osaka University　2-6 Yamadaoka, Suita, 565-0871 Japan

Original manuscript received April 20, 2018, Revisions received April 23, 2020, Accepted June 3, 2020

academic papers in the basic research stage.[14][15]Note 3) Hence, they have not applied the bibliographic clustering method with the analysis of standardization, nor have they discussed the influence of standardization on knowledge accumulation.

Scholars have not considered standardization to be an important factor for innovation for some time. One reason for this is the lack of specific data on standardization. Nevertheless, more recently, empirical analyses targeting the relationship between patents, research papers, and standards have surfaced.[14][15] We discuss the influence of standards to discuss recent research results.

Our major contribution is the discovery that standardization can change the knowledge flow between academic papers and patents with regard to specific technology related to MPEG. Moreover, we identify differences in knowledge creation among certain nations. Thus far, it has been still unclear how standardization activities influenced R&D activities in each nation in terms of bibliographic perspectives.[15]–[17] Finally, we show the scenario to achieve a comprehensive method of achieving more promising technology areas by using elements including academic articles, patents and standards.

2 Literature review

2.1 Overview

Bibliographic analysis of social and management issues related to standardization is still emerging. In prior bibliographic research, there are insufficient scenarios to achieve comprehensive knowledge creation because standards have not been considered as an element of innovation and scholars have not included all the essential elements required for innovation. Hence, the influential estimation of standards has not yet been well established.

Shibata, Kajikawa, and Sakata employed the relationship between academic papers and patent filings for their bibliographic analysis.[18] Nevertheless, they did not use data about standardization. It is largely due to the lack of data about standardization, even in this era of big data.[19][20] Another reason for not employing data about standardization is that, in the existing R&D strategy and national innovation system, policy makers and academia still regard academic papers and patents as the primary output indicators for R&D projects. Thus, there is no positive incentive to measure the influence of standardization. However, in recent national innovation systems of technologically advanced regions such as the US, the EU, and Japan, standardization has become increasingly significant in their innovation policy and corporate strategy. In the EU, standardization has now become an inevitable element for policy evaluation as well as for R&D project evaluation.[21] In the US, a system to evaluate standardization is still under development.[17] In Japan, the government expects the standardization policy to play an essential role.[22] However, currently, it has not been fully implemented. In Japan, as a national project, the New Energy and Industrial Technology Development Organization (NEDO) has employed the results of standardization, albeit only to measure the number of draft proposals.

2.2 Management perspectives

2.2.1 Exploitation and exploration using bibliographic analysis

Bibliographic analysis can help improve the ability of organizations to conduct exploration. It involves less uncertainty in terms of related costs, which have been identified as an obstacle to exploration. Further, it enables organizations to find ways of using the newly discovered knowledge.

Prior research has largely focused on the separative or contradictive relationship between exploration and exploitation. It is said that there should be a balance between exploration and exploitation for organizational learning.

The relationship between the two is not necessarily complementary in practice. This is largely because of the uncertainty when it comes to exploration.[23]–[26] In addition, it is difficult for organizations to evaluate the quality of discovered knowledge because of their low absorption capacity. Hence, even if they obtain external knowledge, they are unable to utilize it.

Importantly, firms need to search their knowledge space for their strategic action. This space comprises an internal sector (within the firms) and an external sector (outside the firms).[27]–[29] In addition to generating internal knowledge, firms can be more innovative in translating the knowledge around them into new products.[30] Therefore, knowledge located outside organizational boundaries plays an important role in firm performance.[31]

Moreover, the breadth of the external search is positively correlated to innovativeness.[32][33] Nevertheless, exploration involves the risk of failure. Therefore, firms need to explore two different dimensions of organizational search, namely, breadth and depth,[25][31][34] but the two dimensions have been found to exhibit a trade-off relationship.

Firms tend to depend on the same internal technologies to produce new products[35][36] and this behavior establishes an environment of path dependency for innovation.[37] It can lead to a competency trap, whereby exploration (i.e., radical innovation) is challenging.[38]

Further, it is difficult to find knowledge space outside a firm's boundary. Organizations are not cognitive of the entire knowledge space. Therefore, the knowledge they find is sometimes incomplete and less than they require, though they consider it to be complete.[39] Hence, their decisions are *bounded rational*.[40]

In addition, firms usually do not create disruptive innovation intentionally. Hence, new concepts are recognized and formed before the research has progressed to a certain stage in public. This implies that finding ways to use these new concepts is also a matter of concern. Timely detection of new technological frontiers brings about a first-mover advantage when it comes to an R&D strategy. The BCCN can help find new technologies in the knowledge space, as well as new ways to use them.[11]

In summary, a lack of information processing technology has, generally, made it difficult for organizations to explore the knowledge space effectively and completely. Hence, exploration and exploitation have been considered as separate approaches, and organizations have traditionally chosen to pursue one or the other.[25] Nevertheless, in terms of R&D strategies, the environment has changed as a result of the rapid expansion of data availability and the development of information processing technology.[19][41] Using the developed bibliographic method, the prior theoretical framework is changed. Applying the BCCN to the relevant documents reduces the cost of searching. Further, it is easy to visualize emerging concepts, which cannot be necessarily described in existing terms. Therefore, we can discover new knowledge space arising from emerging technologies. This is in keeping with the objective to paint a bigger picture of scientific knowledge.[42] Owing to the developments in ICT and big data, exploitation and exploration can now be compatible and complementary. Nevertheless, few studies have investigated this change.

2.3 Methodological perspectives

2.3.1 Science linkage between patents and academic papers

Patents are an important indicator of R&D success and innovation.[43]–[49] Further, the patent citation network contains information about patents and the links between them.[50]–[52] Hence, this is an important source of data for bibliographic analysis. Similar to patents, citations in academic papers also provide important information. Garfield pioneered the use of citation analysis for academic papers.[53] Further, scholars have studied academic paper networks based on co-authorship to analyze knowledge flows.[54] While patents are a private knowledge stream, academic papers are a public knowledge stream.[55] Hence, the notion of science linkage examines the flow of knowledge from public to private entities. This is useful for predicting potential areas of technological development.[18] Thus, various studies have examined citation networks.[56]–[60]

2.3.2 Knowledge space structure and standardization

In addition to the relationship between patents and academic papers, this study uses information on standardization, as shown in Fig. 1. The knowledge space model includes 1) patents, 2) academic papers, and 3) standardization, in contrast to prior research which included only patents and academic papers.[18] As such, organizations can now recognize standardization as a factor in the knowledge space. Further, using this model, we can observe how standards influence the technological similarities between patents and academic papers. In previous studies related to standardization, little attention has been paid to bibliographic analysis.[61][62] However, standardization is now considered important, particularly in fields, such as ICT, and for the specific subject of this study (i.e., an artificial intelligence-related technology).[63]–[65] Thus, we consider standardization as one of the essential components of the knowledge space in this study.

Table 1. Data preparation and processing procedure

Data preparation	Data processing	
	Bibliographic Clustering of Citation Networks (BCCN) [11][18][67]	New additional treatment of this study
We select the relevant academic papers and patents for clustering using the key words: 1) “mpeg” 2) “mpeg” and “standardization” We used the following bibliographic information sources from the Thomson Reuters database for the subtraction process: 1) Title information in Thomson Innovation (patents) 2) Title information in Web of Science (academic papers).	1) Subtracting the patents and academic papers from patent and journal databases using intended key words (this process selects targeted patents and journal papers) 2) Clustering patents and academic papers through unsupervised learning in terms of citation networks (this process clusters patents and papers with similar bibliographic characteristics) (The Academic Landscape* system is used for the computation of clustering and consequent comparison between clusters.)	1) Chart the heat map of cosine similarities between (i) patents and (ii) academic papers (Layer 1). 2) Chart the heat map of cosine similarities between (i) patents and (ii) academic papers (Layer 2). 3) Find the specific technologies whose cosine similarity between patent clusters and academic paper clusters differs between Layer 1 and Layer 2. 4) Observe the science linkage of specified technologies, and find a specific technology region.

Note: * Copyright © 2010-2013 Innovation Policy Research Center, The University of Tokyo, and Copyright © 2012-2013 Kajikawa Laboratory, Graduate School of Innovation Management, Tokyo Institute of Technology

With regard to patents and standardization, prior research has found that standardization activities increase the number of patent applications.[15] Nevertheless, in general, we do not patent standardized technology unconditionally.[66]

2.4 Research questions

We formulate the following research question: Does standardization affect science linkage?

3 Methods and results

3.1 Overview

We employed a data preparation and analysis method, similar to that used in previous studies,[11][18][67] and proposed a procedure involving the integration of exploration and exploitation in the information gathering phase (Table 1 and Appendix Fig. A.1).[67]–[69]

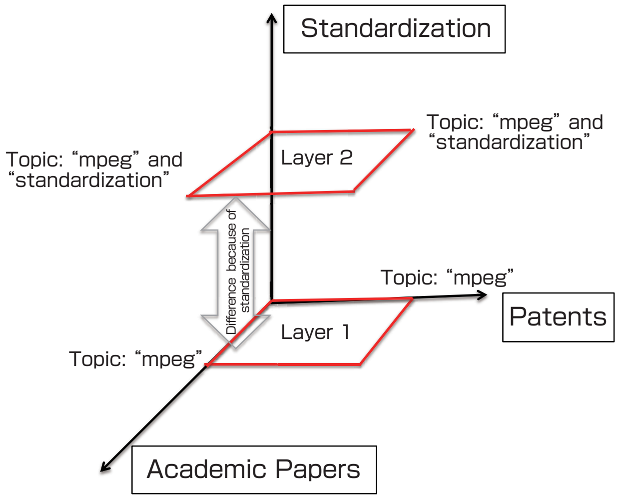


Fig. 1 Three-dimensional knowledge space structure model

3.1.1 Data preparation

We used Web of Science and Thomson Innovation as data sources. The Web of Science is an online database of academic papers, enabling comprehensive citation searches. Thomson Innovation is a global patent database. We used MPEG as the analysis subject. As it is a typical standardized technology, we can clearly observe the influence of standardization on the knowledge space. We extracted published works and patent filings from 1980 to 2014. We connected the keywords “mpeg” and “standardization” with a Boolean operator “AND” and prepared two search strings: 1) (mpeg), and 2) (mpeg AND standardization). Following a keyword search, we selected 6,560 papers and 42,904 patents from the databases for the search string (mpeg), and 1,535 papers and 7,347 patents for the search string (mpeg AND standardization) (See Table 1 and Fig. A.1).[Note 4)]

3.1.2 Research procedure

We used an analysis method similar to that used in previous studies for data preparation; and BCCN.[11][18][67] We also employed an additional procedure in this study (i.e., a new data processing procedure, as listed in Table 1).

3.2 Bibliographic clustering of citation networks

After the clustering computation, 23 paper clusters and 111 patent clusters were obtained for (mpeg) and 14 paper clusters and 39 patent clusters for (mpeg AND standardization) (there was noise among the clusters, because "mpeg" is also used as a scientific phrase in chemistry; however, we ignored these noisy clusters in the analysis).[67] We only used clusters having nodes above 100. The images of the clusters are shown in Figs. 2.1, 2.2, 2.3, and 2.4. The size and major contents of the three largest relevant clusters are given in each figure.Note 5)

3.2.1 Similarity between patent clusters and academic paper clusters

Shibata *et al*. and Iwami *et al*. compared the bibliographic characteristics of clusters of patents and academic papers to study technological similarities and potentially promising technological areas.[18][70] To observe the similarities, we first selected important representative key words (Appendix Table B.1 and Appendix Table B.2) from each cluster using a mutation method (Appendix C) of the term frequency-inverse document frequency (TF-IDF) in Layer 1 and Layer 2. Figures 3 and 4 show the heat maps for Layers 1 and 2, respectively, having calculated the cosine similarities.

3.2.2 Similarities between Layer 1 and Layer 2

Figure 5 shows the method of analyzing the data.[18] The same method is also useful for analyzing discovered knowledge. If the patent is extant and the publication of academic papers is insufficient (Area C), then industrial technology (patents) leads the technology frontier, followed by basic science (academic papers). Hence, we can see the potential for the progressive development of basic research in this area of technology. This implies that for researchers and institutions seeking research themes, the recently obtained information will be highly beneficial for exploitation.

Conversely, when scholars have already published academic papers and if patents are scarce (Area B), the R&D applied to this region is said to not have developed sufficiently; hence, there is a significant opportunity to obtain patents. Thus, the application of this method can help an organization's search efforts in the knowledge space[27]–[29][36][37] and improve its ability to utilize the newly discovered knowledge. Moreover, science linkage is not evident in Areas B and C, whereas it is

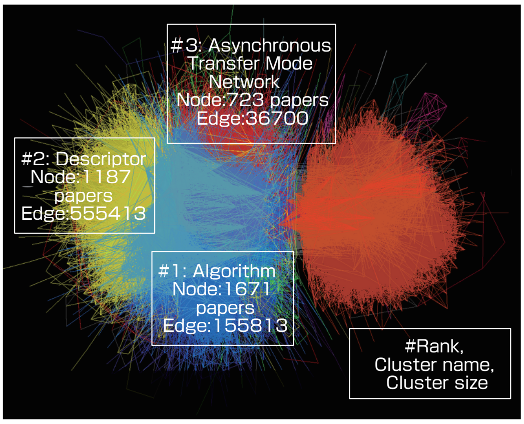


Fig. 2.1 Clusters of academic papers (mpeg)

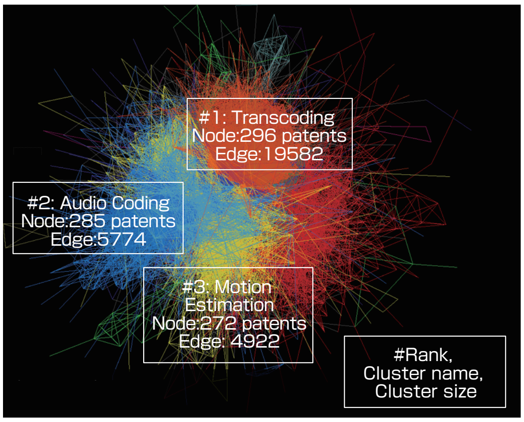


Fig. 2.2 Clusters of academic papers (mpeg AND standardization)

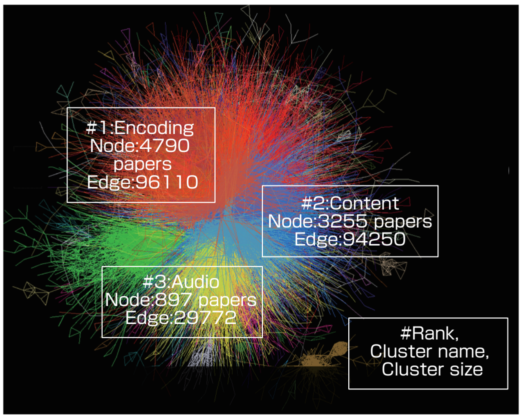


Fig. 2.3 Clusters of patents (mpeg)

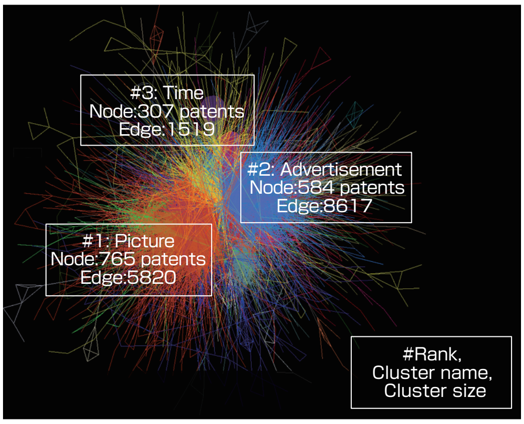


Fig. 2.4 Clusters of patents (mpeg AND standardization)

evident in Area A. If the linkage pattern changes from pattern A to B or from pattern A to C, the science linkage is broken.

The comparison between Layers 1 and 2 is an additional procedure, which has not been included in the method used in prior studies.[18] In Layer 1 of Fig. 1, there is consideration of 1) patents and 2) academic papers, and the key word, as well as our search string, is (mpeg). Similarly, in Layer 2, there is consideration of 1) patents, 2) academic papers, and 3) standardization, and our search string is (mpeg AND standardization). Comparing Layers 1 and 2, we can clearly see the influence of standardization on industrial, technological development (patents) and basic research (academic papers). We use the difference in the heat maps to test the derived hypotheses.

3.2.3 Comparison of national innovation strategy

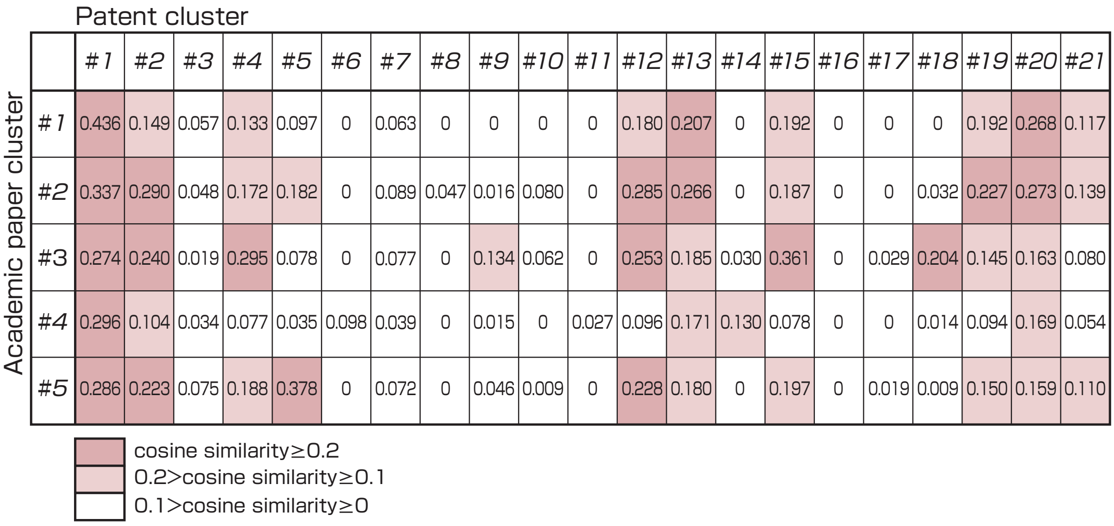
We collect the data related to the country of origin of the patent applicant's institution and the authors of the academic papers. The data are collected in each cluster basis. We use these data to study the differences in the knowledge creation among nations.

4 Discussion

4.1 Overview

Figures 3 and 4 show the cosine similarities between the clusters. The cosine similarity between patent cluster #x and academic paper cluster #y in Layer #z is denoted as follows.

Cosine similarity of Layer #z between patent cluster #x and academic paper cluster #y



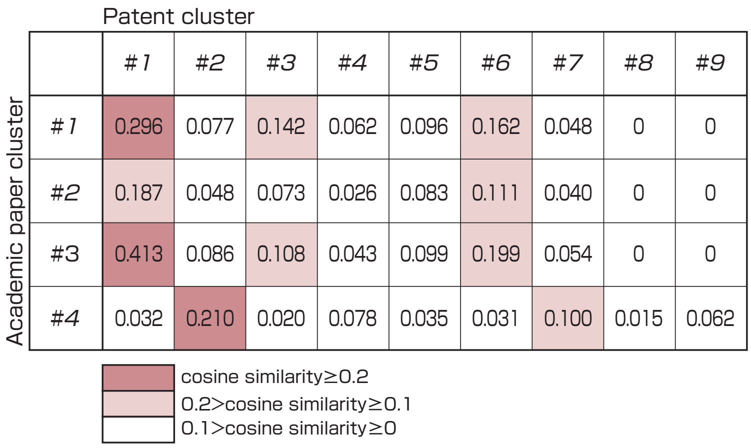
Patent cluster

Academic paper cluster

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21
#1	0.436	0.149	0.057	0.133	0.097	0	0.063	0	0	0	0	0.180	0.207	0	0.192	0	0	0	0.192	0.268	0.117
#2	0.337	0.290	0.048	0.172	0.182	0	0.089	0.047	0.016	0.080	0	0.285	0.266	0	0.187	0	0	0.032	0.227	0.273	0.139
#3	0.274	0.240	0.019	0.295	0.078	0	0.077	0	0.134	0.062	0	0.253	0.185	0.030	0.361	0	0.029	0.204	0.145	0.163	0.080
#4	0.296	0.104	0.034	0.077	0.035	0.098	0.039	0	0.015	0	0.027	0.096	0.171	0.130	0.078	0	0	0.014	0.094	0.169	0.054
#5	0.286	0.223	0.075	0.188	0.378	0	0.072	0	0.046	0.009	0	0.228	0.180	0	0.197	0	0.019	0.009	0.150	0.159	0.110

■ cosine similarity≥0.2
▨ 0.2>cosine similarity≥0.1
□ 0.1>cosine similarity≥0

Fig. 3 Heat map (Layer 1): cosine similarities between patent clusters and academic paper clusters



Patent cluster

Academic paper cluster

	#1	#2	#3	#4	#5	#6	#7	#8	#9
#1	0.296	0.077	0.142	0.062	0.096	0.162	0.048	0	0
#2	0.187	0.048	0.073	0.026	0.083	0.111	0.040	0	0
#3	0.413	0.086	0.108	0.043	0.099	0.199	0.054	0	0
#4	0.032	0.210	0.020	0.078	0.035	0.031	0.100	0.015	0.062

■ cosine similarity≥0.2
▨ 0.2>cosine similarity≥0.1
□ 0.1>cosine similarity≥0

Fig. 4 Heat map (Layer 2): cosine similarities between patent clusters and academic paper clusters (with standardization)

$$= Similarity\ (x, y, z). \qquad (1)$$

We denote the representative key words of each cluster as *Kw(x, y, z)*. When z = 1, it corresponds to Layer 1 and when z = 2, it corresponds to Layer 2. When y = 0, *Kw (x, 0, z)* represents the set of representative key words of the patent clusters in #x of Layer #z. When x = 0, *Kw(0, y, z)* represents the set of representative key words of the academic paper clusters in #y of Layer #z. We assume that there is a strong linkage between factors when the similarity is larger than 0.2.

For example, the cosine similarity in Layer 1 between patent cluster #1 and academic paper cluster #1 is Similarity *(1, 1, 1) = 0.436* (Fig. 3). Tables B1 and B2 list the detected representative key words for each cluster.

4.2 Characteristics of Layer 1

Figure 3 shows the relationship between industry technology (patents) and academic research (papers) under the search string (mpeg). For instance, technologies in the first patent cluster (#1) have terms in common with the academic papers in clusters #1, #2, #3, #4 and #5. We express this as follows.

$$Kw(1,0,1) \cap Kw(0,i,1) \neq \phi, \qquad (2)$$
$$\text{where } i = 1 \text{ to } 5.$$

This is confirmed from the fact that *Similarity (1, j, 1) (j = 1, 2, 3, 4,* and *5)* is greater than 0.2. This reveals that technological information presented in papers is generally patented. Thus, the result of basic research has been industrialized.

Further, among the combinations, *Kw(5,0,1)* (Cluster 5) in the patents and *Kw(0,5,1)* (Cluster 5) in the academic papers in Fig. 3 have common representative key words of "watermark"[Note 6] (Appendix Table B.1). Therefore,

$$Kw(5,0,1) \cap Kw(0,5,1) \neq \phi. \qquad (3)$$

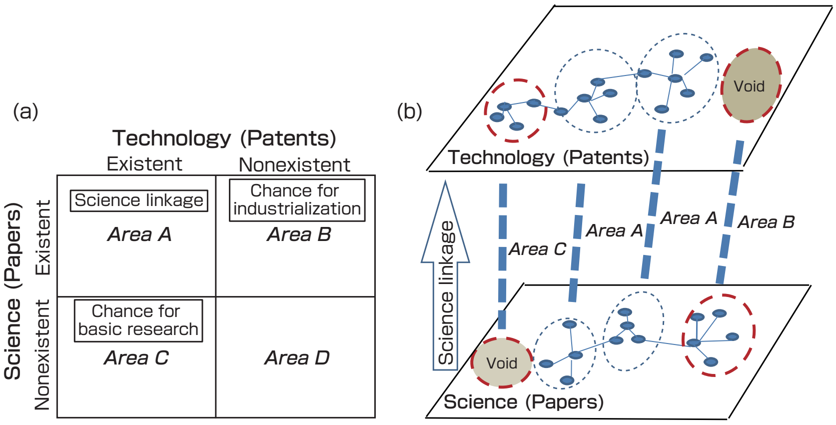
This is supported by the fact that *Similarity (5, 5, 1) = 0.378 > 0.2*. This means that academic research (papers) and industrial applications (patents) advance simultaneously when it comes to watermark technology, and there is a definite science linkage between them.

4.3 Characteristics of Layer 2

Figure 4 shows the relationship between industrial technology (patents) and academic research (papers) using the search string (mpeg AND standardization). The cosine similarity between the two factors is shown. The key words for each cluster are shown (Appendix Table B.2). Patent Cluster 7 has the representative key word "watermark;" however, there is no corresponding academic paper cluster with the term "watermark," as opposed to Layer 1. We denote this relationship as follows.

$$Kw(7,0,2) \cap Kw(0,i,2) = \phi, \qquad (4)$$
$$\text{where } i = 1 \text{ to } 4.$$

Compared to industrial research (patents), basic research results (papers) are scarce when it comes to technologies related to "watermark." Therefore, the science linkage is low. While not hindering science linkages, the standardization



Reference: [18]
Figures are made by the authors.

Fig. 5 Relationship between science and technology: (a) Four types of relationship; (b) Gap between science and technology

process may not improve basic research in the specified technology related to "watermark." In other words, standardization may not necessarily facilitate academic achievement via academic papers. Conversely, the patenting process is not hindered by standardization.

4.4 Difference in knowledge creation among countries

The analysis results of academic papers and patents of each country are shown in Appendix D as well as the tendencies of each country. In cases of China and Korea, the number of academic papers on MPEG-related to standards is higher than that in the case of Japan (Appendix Table D1), whereas the number of patents on MPEG-related to standards is lower or largely the same as that in the case of Japan (Appendix Table D2). This shows that among the East Asian countries, the knowledge creation tendency differs, although the regions are near one another from a geographical perspective (According to the principle of economic geography, economic similarity exists between neighboring regions. This principle, however, cannot be applied to this result.).

On the other hand, in the cases of China and Korea, research is advancing in the area where achievement of basic research is strongly expected. The research target of the two countries is strategic. They find emerging research areas and survey the area, much ahead of other countries. Moreover, in this area, the knowledge creation tendency is similar to that of Germany, whereas that of Japan is different to that of Germany.

The knowledge creation tendency of the US is different from that of other countries. The standardization activity is conducted to obtain patents in the case of the US telecommunication equipment.[72] We assume that the result of this study is in line with the observed facts in the case of the US.

In summary, the knowledge creation of the MPEG-related standard can be categorized as follows: 1) China, Korea, and Germany, 2) Japan, and 3) the US, in Table 2.

4.5 Summary

As for our research question, the results indicate that standardization affects the science linkage between patents. This is supported by the difference between Layers 1 and 2 at the universal level.

Table 2. Summary of the knowledge-creation activities

	Patent	Academic paper	Knowledge creation type
China	Lower	Higher	1
Korea	Lower/Same	Higher	1
Japan	(Baseline)	(Baseline)	2
Germany	Lower	Higher	1
US	Higher	Higher	3

As shown in Fig. 6, the standardization process may hinder academic achievement. In Layer 1, the watermark technology represents clusters in both patents and academic papers, whereas in Layer 2, it represents clusters in patents, but not so in academic papers. This shows that the science linkage has altered between the two layers. The same is confirmed from the fact that *Similarity (5, 5, 1) = 0.378* (Layer 1) decreases to *Similarity (7, 0, 2)*, which ranges from 0 to 0.1 (inclusive) (Layer 2).

We can interpret this result as follows. In certain technologies, such as watermarks, standardization can generally suppress academic achievement. This is in accordance with the results obtained in the Research Center of Material Science in Germany,[15] and is confirmed in this study using the bibliographic method.[Note 7)]

Furthermore, as indicated by the difference between Layers 1 and 2 in Fig. 6, the standardization process may not necessarily hinder patenting. We believe this is because the patented technology does not directly relate to the standardization, but develops around the standardized technologies of MPEG. Therefore, in this case, there is a complementary relationship between standardization and patenting. At the country level, this discussion is in line with the results obtained for Japan and the US, as listed in Table 2.

Figure 7 shows the knowledge structure proposed in this study. The knowledge of standardization is embedded in patent knowledge. Conversely, in academic paper knowledge, the effect of standardization creates a knowledge void. The total academic knowledge is divided into two areas, 1) (mpeg) and 2) (mpeg AND standardization). At the universe level, this means that it is difficult to simultaneously undertake standardization and academic research activities.

The findings of this study indicate that at the regional level. Countries show different knowledge creation tendencies for academic papers and patents related to standards. In particular, Japan is comparably more advanced in obtaining patents related to standards than publishing academic papers, and this result is in line with the previous research result obtained in the US. This result shows that, in terms of the achievements of researchers and engineers, to participate in ICT technology standardization activities is complementary to registering patents.

4.6 Scenario to implement comprehensive knowledge creation

The necessity of the intended scenario to achieve the comprehensive knowledge creation is illustrated in Fig. 8. In the past the knowledge related to academic paper and that related to patent were considered independently and the integration between them did not seem important in the knowledge creation. These two activities are carried out independently. The situation has changed and improved to the next step and the integration between them is considered important now. For example, even in basic research organizations such as universities patent application seem as important as the publication of academic articles.

The scenario of this study implies that the integration between academic articles and patents related to standard is important in the ideal knowledge creation. It is because today's goods and services provide consumers with utility

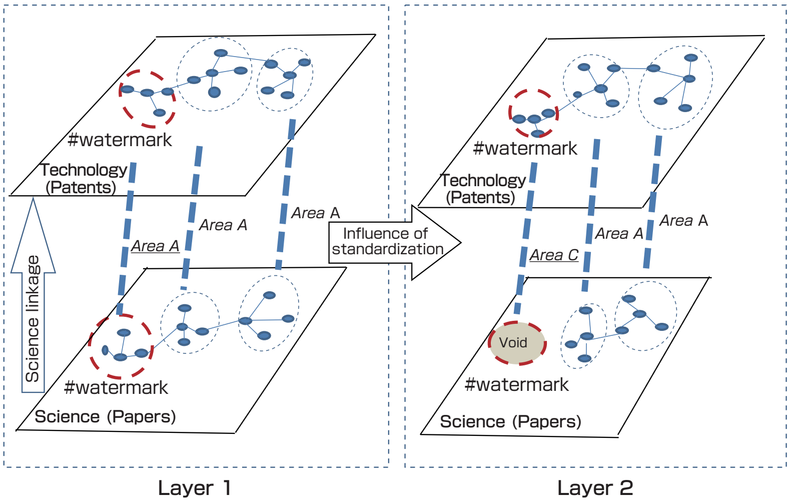


Fig. 6 Change in science linkage arising from standardization

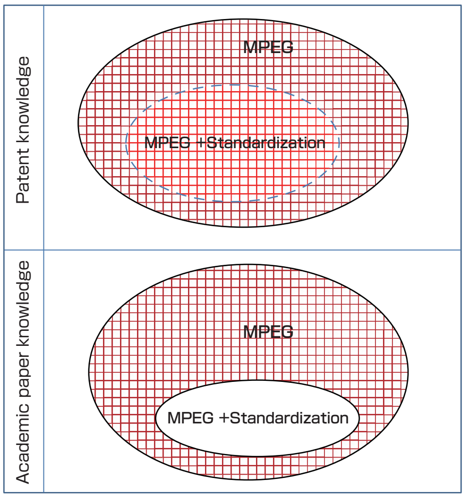


Fig. 7 Corresponding relationship between patents and academic papers in Layer 2

based on network effect. Practically, to achieve net-work effect, goods and services need to use the standardized technology (e.g. MPEG). Figure 8 shows the three types of knowledge creation models and the relationships among academic papers, patents, and standardization for MPEG for different types of countries.

5 Limitations

Our study is not without limitations. We examined the influence of standardization on MPEG technology; however, we should avoid any overgeneralization of these results. To find the same phenomenon elsewhere, it is necessary to identify the mechanism behind the observed results. We conducted a bibliometric analysis of the impact of standards on patents and papers in each specific country. Whether a general underlying cause for the results of our analysis exists remains a question for future research.

Further, this study analyzed pooled data from specific periods. Hence, we cannot observe the dynamics of the relationship between patent cluster formation and academic paper cluster formation. If we had observed the dynamic development of each cluster, we could obtain more valuable information regarding standardization.

6 Conclusion

We examined the influence of image-digitizing technology (MPEG) to consider the ideal knowledge creation under standardization because in MPEG technology, standardization plays an important role. We observed the differences in knowledge creation related to standardization among countries and discuss the characteristics and difference. Based on these results, patents, journal papers, and standards are the important factors to improve innovation as ideal knowledge creation.

This study presents a new knowledge structure model. The structure model uses three-dimensional coordinates (Fig. 1). By comparing the standardization-related and non-standardization-related layers, we addressed the influence of standardization on science linkage. In the case of digital watermark technology, standardization affects the science linkage, and a knowledge flow from patents to academic papers decreases.

Acknowledgements

This study is conducted as a part of the Project at the Research Institute of Economy, Trade and Industry (RIETI) and this research article is based on the Policy discussion paper of RIETI. The author appreciates the valuable comments of Discussion Paper seminar participants at RIETI and the supports from Director Nagano, Director Yamamoto, Director Fukuda and Director Izumi of the Technical Regulations, Standards and Conformity Assessment Unit at the Ministry of Economy, Trade and Industry (METI). This work was supported by JSPS KAKENHI Grant Number 15K03718.

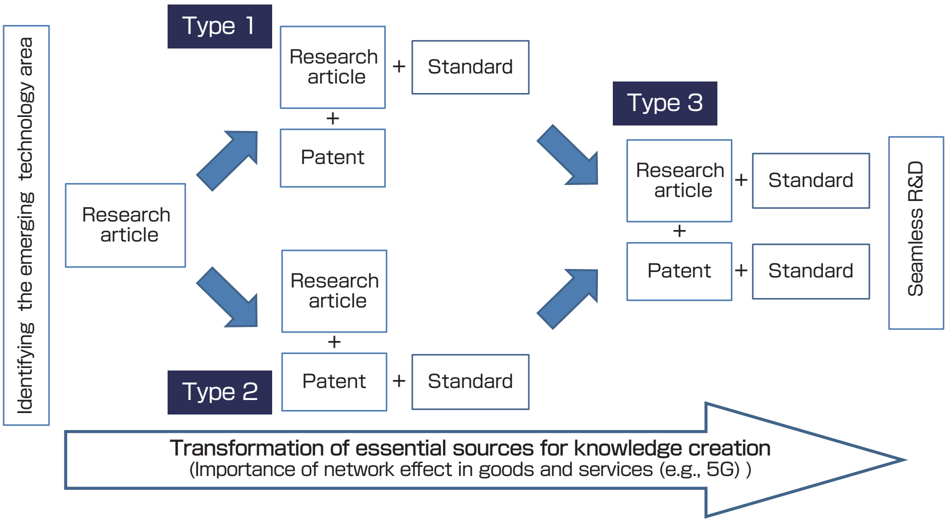
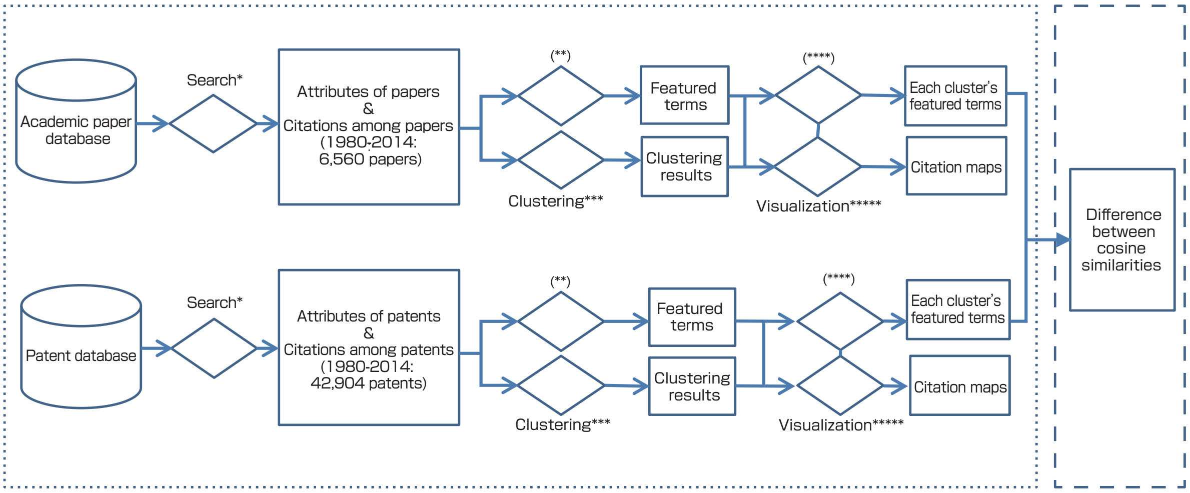


Fig. 8 Scenario for comprehensive knowledge creation

APPENDICES

APPENDIX A:



Note: * In this case, the queries are (1) "mpeg" or (2) "mpeg" and "standardization."
** Natural language processing: NC-Value methods[69].
*** The clustering algorithm[67].
**** Natural language processing: TF-IDF method.
***** The Large Graph Layout (LGL) visualizing engine[68].

Fig. A.1 Detailed flow chart

APPENDIX B:

Table B.1. Keywords in academic paper clusters and patent clusters (Layer 1)

Academic paper cluster	Top TF-IDF terms
#1	coding, video, bit, algorithm, video coding
#2	video, object, image, motion, descriptor
#3	video, traffic, network, atm, error
#4	motion estimation, estimation, search, motion, video
#5	watermarking, video, watermark, quality, watermarking scheme

Patent cluster	Top TF-IDF terms
#1	video, block, frame, encoding, image
#2	content, medium, video, program, user
#3	audio, medium, file, content, player
#4	packet, stream, video, transport, data
#5	content, watermark, digital, file, medium
#6	memory, memory device, memory cell, flash, flash memory
#7	information storage medium e.g, information storage medium, specific unit, information storage, dvd ram
#8	touch, user, touch screen, sensor, electronic
#9	packet, broadcast, digital, stream, data
#10	content, sponsor, communication facility, mobile communication facility, user
#11	card, electronic, case, electronic device, cover
#12	network, audio, image, video, remote
#13	stereoscopic, dimensional, video, image, picture
#14	power, charging, wireless power, power transmission, battery
#15	packet, video, stream, transmission, frame
#16	interferometric, light, interferometric modulator, microelectromechanical, modulator
#17	caption, caption service, transmitting digital broadcast, transmitting digital broadcast signal, closed caption service
#18	wireless, network, communication, node, access
#19	volume descriptor, recording, volume descriptor sequence, descriptor sequence, descriptor
#20	encoding, image, picture, frame, shot
#21	image, light, organic light emitting display, emitting display, light emitting display

Table B.2. Keywords in academic paper clusters and patent clusters (Layer 2)

Academic paper cluster	Top TF-IDF terms
#1	mode, video, h.264/avc, transcoding, rate
#2	audio, mdct, audio coding, transform, dct
#3	object, segmentation, motion, motion estimation, block
#4	descriptor, retrieval, content, multimedia, shape

Patent cluster	Top TF-IDF terms
#1	video, block, picture, motion, frame
#2	content, program, network, video, advertisement
#3	video, packet, recording, stream, time
#4	encrypted, content, key, packet, encryption
#5	dimensional, video, image, picture, stereoscopic
#6	artifact, block, filtering, pixel, video
#7	watermark, content, document, blanking interval, blanking
#8	enhanced data, trellis, vestigial, traffic information data, traffic information
#9	connector, interface, medium, file, card

APPENDIX C: Calculation of Cosine Similarity between Clusters

Using the key words in each cluster, we compared the bibliographic similarities between i) patent clusters and ii) academic paper clusters to reveal the technological analogy. To measure the similarity, we used the cosine similarity formula, expressed as follows.

$$\textit{Cosine similarity}=\frac{\overline{v}1\cdot\overline{v}2}{|\overline{v}1|\cdot|\overline{v}2|}$$

where $\overline{v}1$ and $\overline{v}2$ are vectors of word frequency in each cluster, e.g., $\overline{v}1$ = (frequency of word 1, frequency of word 2,…, frequency of word n) = (f_1, f_2,…, f_n).

The cosine similarity ranges from 0 to 1. When the similarity is 1, the two vectors of the cluster are identical, and the two clusters are similar. Conversely, when the similarity is 0, the two vectors of the clusters are completely different sets of words, and therefore, the two clusters are different. We depict the calculated values in the heat map, where a deep color represents a high degree of similarity whereas the white areas indicate the absence of a relationship. The heat map shows 1) the most related clusters and 2) the most unrelated clusters.

APPENDIX D:

Table D.1. Activities of countries: Academic papers (MPEG related to standards)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
China	4.500	1.909	1.308	3.000
Korea	4.700	2.545	1.769	6.000
Japan	1.000	1.000	1.000	1.000
Germany	3.300	1.909	2.077	4.800
US	7.200	6.545	5.000	8.200

Note: The figure is the ratio of the number of academic papers of each country to Japan. Underlined numbers are over 1.

Table D.2. Activities of countries: Patents (MPEG related to standards)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
China	0.191	0.025	0.063	0.100	0.118	0.500	0.333
Korea	0.723	0.350	0.281	1.200	1.235	1.167	0.833
Japan	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Germany	0.064	0.125	0.063	0.400	0.000	0.167	0.000
US	3.340	4.375	1.313	4.400	1.647	4.167	5.167

Note: The figure is the ratio of the number of patents of each country to Japan. Underlined numbers are over 1.

Table D.3. Activities of countries: Academic papers (MPEG)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
China	10.288	2.172	2.514	1.195	0.413
Korea	2.603	2.828	4.054	1.610	0.810
Japan	1.000	1.000	1.000	1.000	1.000
Germany	0.822	1.108	1.946	0.415	0.381
US	3.603	4.441	6.081	4.902	1.381

Note: The figure is the ratio of the number of academic papers of each country to Japan. Underlined numbers are over 1.

Table D.4. Activities of countries: Patents (MPEG)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
China	0.113	0.131	0.744	0.143	0.246	0.462	0.016
Korea	0.628	1.069	3.256	0.643	1.404	12.846	0.170
Japan	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Germany	0.079	0.150	0.674	0.092	0.158	0.462	0.004
US	1.919	6.100	13.488	3.622	4.456	16.769	0.075

Note: The figure is the ratio of the number of patents of each country to Japan. Underlined numbers are over 1.

Notes

Note 1) MPEG is the name of the standardized technology given by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC).[1] MPEG is the abbreviation of Motion Picture Experts Group, which is a working group of authorities involved in the preparation of the technology standard. Later, the expert group name was adopted as the name of the standardized technology. Hence, in this study, MPEG refers to the technology, which is used to encode, decode, and subsequently transmit video and audio files. Today, MPEG is widely used to send videos over the Internet.

Note 2) “Science linkage” generally means the knowledge flow form academic papers to patents.

Note 3) This study uses a different wording from that of the OECD Frascati Manual. The terms “basic research” and “applied research” follow the definitions in the previous study.[18] We continue this analysis with the definitions and framework that was applied in the prior research. We define “basic research” as academic article research and “applied research” as patent acquisition.

Note 4) The upper flowchart path in Fig. A.1. shows the analysis applied to academic papers. The lower flowchart path shows the analysis applied to the patent literature. The last process compares the results of these two analyses. As noted in Fig. A.1, this method is similar to that commonly applied in previous studies. An introduction to algorithms is outside the scope of this paper, and interested readers are referred to the cited literature.

Note 5) The annotation of these figures follows the descriptions of standard graph theory. The figures also show three large nodes to provide an overview of the structure of the bibliographic sources of academic papers and patents. Graph theory defines a node as a nodal point or endpoint and an edge as the line that connects the nodes. Each cluster is made up of edges and nodes. In Fig. 2.1, the numbers #1, #2, and #3 indicate the order of the size of the nodes (#Rank) within the clusters and correspond to the clusters in Table B.1. Fig. 2.1 illustrates the three largest clusters from Table 2.1. Table B.1, on the other hand, shows all clusters above a certain size. The information about the size of each cluster is shown as the number of nodes and edges (Cluster Size). In Fig. 2.1, #1 is the largest node with 1671 papers, #2 is the next largest with 1187 papers, and #3 is the next largest with 723 papers. The same rule is used for the rest of Fig 2. In addition, a representative extraction word (cluster name) is described to represent the characteristics of each node.

Note 6) “Watermark” here implies a digital watermark technology whereby invisible signals are incorporated into imaging data. This is used for authenticity validation, copyright tracking of digitized imaging, and detections of copyright infringement.[71] It differs from conventional watermarks, which are translucent marks on prints or pictures.

Note 7) At the country level, Germany has a higher academic paper production number than Japan. This is a result of a comparison between the two countries; hence, it does not necessarily mean that academic paper production is improved because of standardization in Germany. Therefore, the discussion here does not have a discrepancy.

References

[1] H. Yasuda: Standardization activities on multimedia coding in ISO, *Signal Processing: Image Communication*, 1 (1), 3–16 (1989).

[2] J. Albusac, J. J. Castro-Schez, L. M. Lopez-Lopez, D. Vallejo and L. Jimenez-Linares: A supervised learning approach to automate the acquisition of knowledge in surveillance systems, *Signal Processing*, 89 (12), 2400–2414 (2009).

[3] G. Amato, P. Savino and B. Magionami: Use of weighted visual terms and machine learning techniques for image content recognition relying on mpeg-7 visual descriptors, *MS '08 Proceedings of the 2nd ACM workshop on multimedia semantics*, 60–63 (2008).

[4] G. Fernández, H. Kalva, P. Cuenca and L. Orozco-Barbosa: A first approach to speeding-up the inter mode selection in MPEG-2/H.264 transcoders using machine learning, *Multimedia Tools and Applications*, 35 (2), 225–240 (2007).

[5] G. Fernández-Escribano, H. Kalva, P. Cuenca and L. Orozco-Barbosa: Very low complexity MPEG-2 to H.264 transcoding using machine learning, *Proceeding of the ACM Multimedia 2006*, 931–940 (2006).

[6] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, L. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel and D. Hassabis: Mastering the game of Go with deep neural networks and tree search, *Nature*, 529, 484–489 (2016).

[7] A. Colbert, N. Yee and G. George: The digital workforce and the workplace of the future, *Academy of Management Journal*, 59 (3), 731–739 (2016).

[8] E. Felten: Preparing for the future of artificial intelligence, The White House, https//www.whitehouse.gov/blog/2016/05/03/preparing-future-artificial-intelligence, Accessed 2016-07-01.

[9] L. Leydesdorff: *The Challenge of Scientometrics: The Development, Measurement, and Self-Organization of Scientific Communications*, DSWO Press, Amsterdam (1995).

[10] N. Shibata, Y. Kajikawa, Y. Takeda and K. Matsushima: Detecting emerging research fronts based on topological measures in citation networks of scientific publications, *Technovation*, 28 (11), 758–775 (2008).

[11] A. Tashiro, H. Tashiro, S. Iwami and I. Sakata: Bibliometric analysis of net zero energy building, *Proceedings of the PICMET '13*, (2013).

[12] W. M. Cohen, R. P. Nelson and J. P. Walsh: Links and impacts: The influence of public research on industrial R&D, *Management Science*, 48 (1), 1–23 (2002).

[13] B.V. Looy, E. Zimmermann, R. Veugelers, A. Verbeek, J. Mello and K. Debackere: Do science-technology interactions pay off when developing technology? An exploratory investigation of 10 science-intensive technology domains, *Scientometrics*, 57 (3), 355–367 (2003).

[14] K. Blind and S. Gauch: Research and standardisation in nanotechnology: Evidence from Germany, *The Journal of Technology Transfer*, 34 (3), 320–342 (2009).

[15] A. Zi, and K. Blind: Researchers' participation in standardisation: A case study from a public research institute in Germany, *The Journal of Technology Transfer*, 40 (2), 346–360 (2015).

[16] S. Tamura: Effects of integrating patents and standards on intellectual property management and corporate innovativeness in Japanese electric machine corporations, *International Journal of Technology Management*, 59 (3/4), 180–202 (2012).

[17] G. Tassey: Method for assessing the economic impacts of government R&D, National Institute of Standards & Technology (2003).

[18] N. Shibata, Y. Kajikawa and I. Sakata: Extracting the commercialization gap between science and technology—Case study of a solar cell, *Technological Forecasting and Social Change*, 77 (7), 1147–1155 (2010).

[19] G. George, M. R. Haas and A. Pentland: Big data and management, *Academy of Management Journal*, 57 (2), 321–326 (2014).

[20] S. Tamura: Generic definition of standardization and the correlation between innovation and standardization in corporate intellectual property activities, *Science and Public*

Policy, 40 (2), 143–156 (2013).
[21] J. Edler, L. Georghiou, K. Blind and E. Uyarra: Evaluating the demand side: New challenges for evaluation. *Research Evaluation*, 21 (1), 33–47 (2012).
[22] Cabinet Office: *The Fifth Science and Technology Basic Plan*, (2016).
[23] C. A. O'Reilly and M. L. Tushman: Ambidextrous organizations: Managing evolutionary and revolutionary change, *California Management Review*, 38 (4), 8–29 (1996).
[24] C. A. O'Reilly and M. L. Tushman: The ambidextrous organization, *Harvard Business Review*, 82 (4),74–81 (2004).
[25] J. G. March: Exploration and exploitation in organizational learning, *Organization Science*, 2 (1), 71–87 (1991).
[26] Z. L. He and P. K. Wong: Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis, *Organization Science*, 15 (4), 481–494 (2004).
[27] R. M. Grant: Prospering in dynamically-competitive environments: Organizational capability as knowledge integration, *Organization Science*, 7 (4), 375–387 (1996).
[28] E. Mansfield: The speed and cost of industrial innovation in Japan and the United States: External vs. internal technology, *Management Science*, 34 (10), 1157–1168 (1988).
[29] L. Rosenkopf and A. Nerkar: Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry, *Strategic Management Journal*, 22 (4), 287–306 (2001).
[30] R. Katila: New product search over time: Past ideas in their prime?, *Academy of Management Journal*, 45 (5), 995–1010 (2002).
[31] K. Laursen and A. Salter: Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms, *Strategic Management Journal*, 27 (2), 131–150 (2006).
[32] L. Dahlander, S. O'Mahony and D. M. Gann: One foot in, one foot out: How does individuals' external search breadth affect innovation outcomes?, *Strategic Management Journal*, 37 (2), 280–302 (2016).
[33] M. Tortoriello: The social underpinnings of absorptive capacity: The moderating effects of structural holes on innovation generation based on external knowledge, *Strategic Management Journal*, 36 (4), 586–597 (2015).
[34] R. Katila and G. Ahuja: Something old, something new: A longitudinal study of search behavior and new product introduction, *Academy of Management Journal*, 45 (6), 1183–1194 (2002).
[35] C. E. Helfat: Evolutionary trajectories in petroleum firm R&D, *Management Science*, 40 (12), 1720–1747 (1994).
[36] J. Wade: A community-level analysis of sources and rates of technological variation in the microprocessor market, *Academy of Management Journal*, 39 (5), 1218–1244 (1996).
[37] R. R. Nelson and S. G. Winter: *An Evolutionary Theory of Economic Change*, Harvard University Press, Cambridge, MA (1982).
[38] B. Levitt and J. G. March: Organizational learning, *Annual Review of Sociology*, 14, 319–340 (1988).
[39] D. Kahneman and A. Tversky: Prospect theory: An analysis of decision under risk, *Econometrica*, 47 (2), 263–291 (1979).
[40] H. A. Simon: Rational decision making in business organizations, *American Economic Review*, 69 (4), 493–513 (1979).
[41] G. George, E. C. Osinga, D. Lavie and B. A. Scott: Big data and data science methods for management research, *Academy of Management Journal*, 59 (5),1493–1507 (2016).
[42] K. Borner, C. Chen and K. W. Boyack: Visualizing knowledge domains, *Annual Review of Information Science and Technology*, 37 (1), 179–255 (2003).
[43] Z. J. Acs and D. B. Audresch: Patents as a measure of innovative activity, *Kyklos*, 42 (2), 171–180 (1989).
[44] Z. Griliches (ed): *R&D, Patents, and Productivity*, University of Chicago Press, Chicago (1984).
[45] Z. Griliches: Patent statistics as economic indicators: A survey, *Journal of Economic Literature*, 28 (4), 1661–1707 (1990).
[46] A. B. Jaffe: Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value, *American Economic Review*, 76 (5), 984–1001 (1986).
[47] A. J. Salter and B. R. Martin: The economic benefits of publicly funded basic research: A critical review, *Research Policy*, 30 (3), 509–532 (2001).
[48] M. Trajtenberg: A penny for your quotes: Patent citations and the value of innovations, *The RAND Journal of Economics*, 21 (1), 172–187 (1990).
[49] M. Trajtenberg: *Economic Analysis of Product Innovation: The Case of CT Scanners*, Harvard University Press, Cambridge, MA (1990).
[50] M. P. Carpenter and F. Narin: Validation study: Patent citations as indicators of science and foreign dependence, *World Patent Information*, 5 (3), 180–185 (1983).
[51] M. P. Carpenter, F. Narin and P. Woolf: Citation rates to technologically important patents, *World Patent Information*, 3 (4), 160–163 (1981).
[52] M. Trajtenberg, R. Henderson and A. Jaffe: University versus corporate patents: A window on the basicness of invention, *Economics of Innovation and New Technology*, 5 (1), 19–50 (1997).
[53] E. Garfield: Citation indexes for science: A new dimension in documentation through association of ideas, *Science*, 122, 108–111 (1955).
[54] I. Demirkan, D. L. Deeds and S. Demirkan: Exploring the role of network characteristics, knowledge quality, and inertia on the evolution of scientific networks, *Journal of Management*, 39 (6), 1462–1489 (2013).
[55] K. G. Huang and F. E. Murray: Does patent strategy shape the long-run supply of public knowledge? Evidence from human genetics, *Academy of Management Journal*, 52 (6), 1193–1221 (2009).
[56] I. Demirkan and S. Demirkan: Network characteristics and patenting in biotechnology: 1990–2006, *Journal of Management*, 38 (6), 1892–1927 (2012).
[57] G. S. McMillan, F. Narin and D. L. Deeds: An analysis of the critical role of public science in innovation: The case of biotechnology, *Research Policy*, 29 (1), 1–8 (2000).
[58] A. L. Oliver: Biotechnology entrepreneurial scientists and their collaborations, *Research Policy*, 33 (4), 583–597 (2004).
[59] A. L. Oliver and J. P. Liebeskind: Three levels of networking for sourcing intellectual capital in biotechnology, *International Studies in Management & Organization*, 27 (4), 76–103 (1997).
[60] L. G. Zucker, M. R. Darby and J. S. Armstrong: Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology, *Management Science*, 48 (1), 138–153 (2002).
[61] W. B. Arthur: Competing technologies, increasing returns, and lock-in by historical events, *Economic Journal*, 99 (394), 116–131 (1989).
[62] P. A. David: Clio and the economics of QWERTY, *American Economic Review*, 75 (2), 332–337 (1985).
[63] T. M. Egyedi and M. H. Sherif: Standards dynamics through an innovation lens: Next-generation ethernet networks, *IEEE Communications Magazine*, 48 (10), 166–171 (2010).

[64] K. Jakobs, R. Procter and R. Williams: Standardisation, innovation and implementation of information technology, *Computers and Networks in the Age of Globalization*, 57, 201–217 (2001).

[65] M. H. Sherif: A framework for standardization in telecommunications and information technology, *IEEE Communications Magazine*, 39 (4), 94–100 (2001).

[66] S. Tamura: A new intellectual property metric for standardization activities, *Technovation*, 48–49, 87–98 (2016).

[67] M. E. J. Newman: Fast algorithm for detecting community structure in networks, *Physical Review E*, 69 (6),066133 (2004).

[68] A. T. Adai, S. V. Date, S. Wieland and E. M. Marcotte: LGL: Creating a map of protein function with an algorithm for visualizing very large biological networks, *Journal of Molecular Biology*, 340 (1), 179–190 (2004).

[69] K. Frantzi, S. Ananiadou and H. Mima: Automatic recognition of multi-word terms: The C-value /NC-value method, *International Journal on Digital Libraries*, 3 (2), 115–130 (2000).

[70] S. Iwami, K. Kogo, F. Tacoa, J. Mori, Y. Kajikawa and I. Sakata: Cross-disciplinary methodology for detection of collaborative and competitive candidates, *Proceedings of the 24th International Conference for Management of Technology (IAMOT 2015)*, South Africa (2015).

[71] D. Choi, H. Do, H. Choi and T. Kim: A blind MPEG-2 video watermarking robust to camcorder recording, *Signal Processing*, 90 (4),1327–1332 (2010).

[72] N. Gandal, N. Gantman and D. Genesove: Intellectual property and standardization committee participation in the US modem industry, in S. Greenstein and V. Stango (eds.), *Standards and Public Policy*, Cambridge University Press, New York, 208–230 (2007).

Authors

TAMURA Suguru

Joined the Ministry of International Trade and Industry, and worked at the Technology Evaluation and Research Division, Industrial Science and Technology Policy and Environment Bureau, and at the Council for Science and Technology Policy, Cabinet Office. Main accomplishments include publication in the *Synthesiology*, *Journal of the American Ceramic Society* and others. He was in charge of research planning and the correspondence author of this article in this study.

IWAMI Shino

IWAMI Shino was a specially appointed researcher in the Institute of Laser Engineering, Osaka University in the fields of data science in relation to global strategy and image recognition. Her specialties are technology management, cloud computing, ubiquitous computing, cyber law, and mechanical and space engineering. She was in charge of data processing in this study.

SAKATA Ichiro

SAKATA Ichiro is a professor at the Graduate School of Engineering and Institute for Future Initiatives, the University of Tokyo (UTokyo). He has held several appointments in UTokyo including the Special Advisor to the President. His research interests include innovation management and policy. He was in charge of the comprehensive supervision in this study.

Discussions with Reviewers

1 Overall

Comment (KOBAYASHI Naoto, Waseda University)

This study shows very suggestive research results which are related to the impact of standardization to the science linkages and to the comprehensive knowledge formation, by examining the relationship between research papers, patents, and standardization for MPEG technology. The results indicate that, among other things, (1) conducting standardization activities and academic research at the same time is difficult and (2) conducting standardization activities and patenting at the same time is possible to a certain degree. An international comparison reveals that the relationship between the three subjects is unique to each country. Based on these findings, a scenario of the dynamics of knowledge creation is presented.

Answer (TAMURA Suguru)

A scenario has been added to the revised version of the manuscript. We not only present scenarios, but we extracted tens of thousands scientific papers. The results of bibliometric analysis are presented.

2 Impact of standardization on science linkages

Comment (KOBAYASHI Naoto)

In this analysis, the focus is placed on the technical term, *watermark*, in particular. The results reveal that there is little relationship between research papers and standardization while a relationship exists between patents and standardization. However, as shown in *Kw*(1,1,2), *Kw*(1,3,2), and *Kw*(2,4,2), there are keywords such as *video* and *content*. This implies that standardization seems not necessarily to hinder the science linkage.

Answer (TAMURA Suguru)

Watermark technology, which is an important part of MPEG technology, is an appropriate subject for analyzing the effects of standardization. Thus, in the conclusions of this paper, the effects of standardization on a specific technology area (watermark) are discussed in the conclusion. In terms of other general words and phrases, it is difficult to state a causal relationship with standardization.

3 Scenario

Comment (KOBAYASHI Naoto)

This study has revealed, based on the evidence, that the science linkage between research papers and patents changes due to the relation to standardization. Based on these results, authors present a scenario of a methodological proposal for what the relationship among research papers, patents and standardization should be. The three subjects described above contribute to the creation of knowledge (Exploration), the use of technology (Exploitation), and its dissemination (Dissemination), respectively. They are considered to make an overall contribution to knowledge formation through their roles and mutual interactions.

Comment (YUMOTO Noboru, National Cerebral and Cardiovascular Center)

There are three ways in which standards are involved in the analysis of each country's knowledge creation activities in the scenario: Type 1, where the standard is linked to the paper. Type 2, where the standard is linked to the patent. Type 3, which can be considered the most advanced of the three, where the standard is linked to the paper and the patent. If Type 3 is not the most advanced form, then a parallel diagram of all three types may be possible. In any case, the results of the analysis of each country are one of the main points of this study. Thus, it is important to draw a diagram by considering the relationship between the scenario and the results of analysis of each country.

Answer (TAMURA Suguru)

For the scenario, we present the temporal development pattern of the three types of relationships between research papers, patents, and standardization as an explanation of the development process of knowledge creation. In addition, we have shown the relationship of the three types to the pattern of countries.