A link is established between the two documents in the event that these indices exceed the threshold value. The network created changes acutely depending on which of the above indices are selected and how the threshold value is established. As the aim of this research is to investigate the extent to which top-down type of classes are reflected in links in document content that is sought bottom-up, when forming networks, we selected the index and determined the threshold values so as to reflect most the given classes.

4.3 Network Clustering

In this research, other than given classes, we conducted clustering using the Newman method in order to carry out labeling of each document based on networks. The Newman method is a widely used method for clustering networks (Newman, 2004). It can be applied even if the number of clusters is unknown; in recent years, it is also being widely applied in large-scale network analysis, such as SNS and blog, due to it being scalable in regard to increases in node numbers. As shown in formula (7), clustering was conducted by maximizing modularity Q, an index for evaluating the modularity of defined networks.

$$Q = \sum_i (e_{ii} - a_i^2) \qquad (7)$$

Here, element e_{ij} of line e represents the fraction of the total number of edges of the number of edges that connect cluster i with cluster j, and a_i represents the sum of row i of line e. Maximizing Q corresponds with maximizing the disparity between the number of edges that exist within clusters and the number of edges that link clusters together.

5. RESULTS

5.1 Extraction of Characteristic Words

We conducted the extraction of characteristic words using the methods outlined in section 3. Table. 1 and 2 show the number of occurrences by category for the 10 words in which tfidf is at the top. We see from these tables that all characteristic words extracted here occur selectively in each class. In addition, when focusing on the combination of class and occurring word, we see that words that aptly reflect the characteristics of the class are being extracted, such as "injected part" occurring largely in the diapedesis class and "allergia" occurring largely in the drug sensitivity class. On the other hand, as the index in this research uses tfidf as an index of the degree of characteristics, general words do not appear at the top, even if the frequency of occurrence is high. For example, although "patient" had the highest number of occurrences, as this word appears in most case studies, tfidf as an index of class capability of documents takes on a low value and does not appear at the top.

Table.1characteristic word and class of operation

characteristic word	diapedesis	flow rate	object person	forget to dose	composition
insulin	0	0	4	0	0
furosemide	0	1	0	0	0
injected part	14	0	9	0	1
allergia	0	0	5	0	0
setup	0	51	0	0	0
flow rate	1	38	2	0	0
coinjection	0	3	7	0	3
anticancer	3	3	7	0	0
periphery	1	2	10	0	2
set	0	5	1	13	1

characteristic word	regimen	name of drug	amount of drug	drug sensitivity	amount and regimen
insulin	0	22	16	0	0
furosemide	4	13	17	0	0
injected part	0	0	1	0	0
allergia	0	1	0	19	0
setup	0	0	4	0	0
flow rate	2	0	9	0	0
coinjection	4	22	9	1	0
anticancer	0	2	5	0	7
periphery	4	4	2	3	0
set	1	9	16	0	6

Table2. Characteristic word and class of treatment

characteristic word	chemo treatment	contraindi-cated drug	dowry of drugs	preparation of drugs
insulin	0	0	3	2
furosemide	0	2	10	1
injected part	14	1	0	0
allergia	0	23	0	0
setup	14	0	0	0
flow rate	8	0	0	0
coinjection	5	7	0	0
anticancer	26	0	2	0
periphery	5	5	0	0
set	0	1	29	3

5.2 Network Analysis

In this research, we aim at uncovering links between classes granted top-down and clusters discovered bottom-up. Thus, we created a network so that documents belonging to identical classes relating to the identical treatment and operation are grouped together the most. To achieve this, we used the index of Class Closeness (CC) defined in formula (8).

$$\mathrm{CC} = \sum_i (\mathrm{D_{ii}} - \overline{\mathrm{D_i}}) \qquad (8)$$

Here, D_{ij} of line D calculates the distance from all of the nodes within category i to all of the nodes within category j and takes the averages of these. Also, $\overline{D_i}$ is the average of the i row of D. The classes here refer to 32 cross classes that cross calculate the class of treatment and the class of operation. CC taking a high value indicates that the nodes of identical classes come close to cases seen across the entire network. If the threshold value of the co-occurrence index is high, there is a tendency for CC to become high as only links with strong links remain. On the other hand, if the threshold value is made too high, a large proportion of the links are lost, the maximum number of connections (LC) of nodes within the network decreases, and the analysis of overall links becomes impossible. Thus, in this research, in order that the product of the maximum number of connections of CC and nodes (CCLC) are at the maximum, we conducted the selection of co-occurrence indices and the determination of threshold values. Table 3 shows the maximum value of CCLC for each co-occurrence index and the values of indices in these instances. By doing so, we discovered that a network that reflects given classes could be obtained by using the Cos coefficient for networking the content and the Jaccard coefficient for networking the background and solutions. Although the Cos coefficient and Jaccard coefficient demonstrate almost identical CCLC in the abstract, background, and solution, it is markedly low regarding the Simpson coefficient. This is because documents that include large numbers of characteristic words are connected with many documents and are close to documents that belong to other classes in the network.

Table 3. Class closeness and co-occurrence index

abstract

Index	Value	CC	LC	CCLC
jaccard	0.290	3.162	510	1612.4
cos	0.459	2.908	594	1727.9
simpson	1.000	0.430	750	322.5

back ground

Index	Value	CC	LC	CCLC
jaccard	0.334	1.441	580	835.8
cos	0.580	3.348	241	806.8
simpson	0.860	0.192	970	186.5

solution

Index	Value	CC	LC	CCLC
jaccard	0.338	1.441	550	1444.7
cos	0.544	2.955	428	1264.9
simpson	0.700	0.279	1023	285.3

Table 4.cluster and characteristic class

cluster index	characteristic class	share of the class	ratio of the class in the cluster	number of the node
1	dowry of drugs	84.80%	60.50%	147
2	preparation of drugs	58.80%	51.30%	111
3	flow rate	55.30%	70.30%	37
4	diapedesis	55%	30%	27
5	-	-	-	25

6. DISCUSSION

The characteristic words displayed in Table 1 and 2 were all selected based on actual data. Networks obtained from similarities with these characteristic words remove the effects of similarities shared in common with all documents and are formed from the overall combination of an independent degree of similarity between two documents. This kind of network is first realized by extracting characteristic words bottom-up. In cases where keywords that should be checked top-down are decided, there are instances when, after having conducted class, there is no guarantee that that keyword is not valid and a network of documents linked only by the characteristic similarities such as those described above cannot be obtained.

Figure 2 shows the abstract network from among the networks created based on the characteristic words obtained bottom-up. In addition, Table 4 shows the abstract network resulting from having conducted clustering using the Newman method. Table 4 shows the top five clusters of the number of nodes and the main classes within these. The clusters from position 1 to 4 clearly reflect the granted classes. In the cluster in position 2, the class of treatment is prioritized and collects together documents classified as drowsy of drugs and preparation of drugs; in positions 3 and 4, the class of operation is prioritized and clusters reflecting flow rate and diapedesis are created; however, in position 5, classes that clearly characterize clusters could not be seen.

Moreover, regarding the solution network, links between clusters and classes such as those described above were extracted. On the other hand, clusters that clearly reflected the given class could not be seen among the clusters of the background network. The fact that networks created through background do not reflect overall given classes is in accord with the results of analysis in which the value of CC corresponding to the co-occurrence index used in creating networks in Table 3 is around 3 in abstract and solution, while in background it is less than 1.5.

So, is the class of background not related in some way to networks created bottom-up? In order to investigate this, we conducted an investigation into how many identical classes exist within the k step from certain nodes. The results of this investigation are shown in Fig. 3. We can see that in no matter which network, the ratio of nodes with identical class decreases to the extent of the number of steps increasing. Here, attention should be drawn to the point that the proportion of identical classes existing within a single step is almost the same in each of the categories of abstract, background, and solution. This indicates that the same classification also takes place in case studies with an extremely high degree of similarity in background. On the other hand, when looking at the overall network, given classes relating to the background are scattered compared with abstract and solution. This concept is outlined in Fig. 4.

With regard to the abstract and solution, links determined from ontology reflect comparatively given categories; however, regarding the background, when looked from the perspective of ontology, there is the possibility that it is completely different from the classes of abstract and solution, even those matters in proximity to step numbers, i.e., it is difficult to predict accidents that occur from information on background. Therefore, it shows that the content of descriptions of background is inadequate from the perspective of preventing reoccurrences. On the other hand, regardless of differences between abstract, background, and solution, the reporter is likely to have taken care to use descriptions that express accurately the maximum limits and the reality of the situation. Class methods that give appropriate suggestions to on-site reporters are required. We can say that there is a strong desire for an ontological construction relating to descriptions of background.

Considering the similarities between only two documents without using network analysis is equivalent to looking at the similarities of the nodes within step one in Fig. 3. In this case, in order to obtain almost identical values between abstract, background, and solution, differences relating to the power of expression of descriptions in the three divisions are not seen. The discovery that descriptions of background are inadequate was made possible for the first time by considering the overall similarities through the networking of case studies.

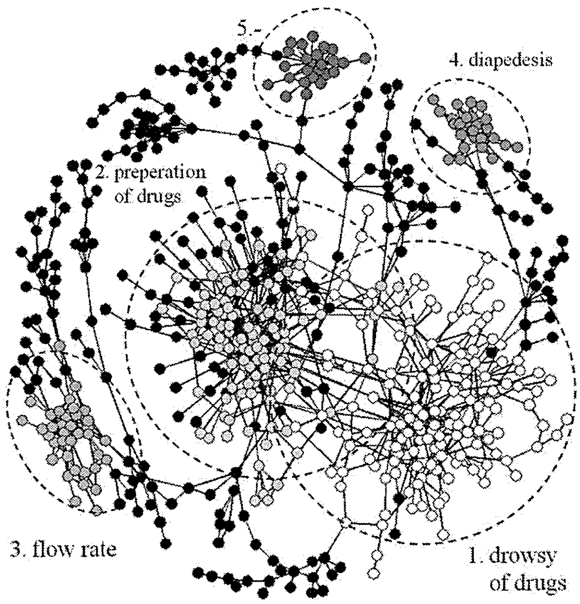

Figure 2. Network of an abstract in Near-miss/ Averse Event

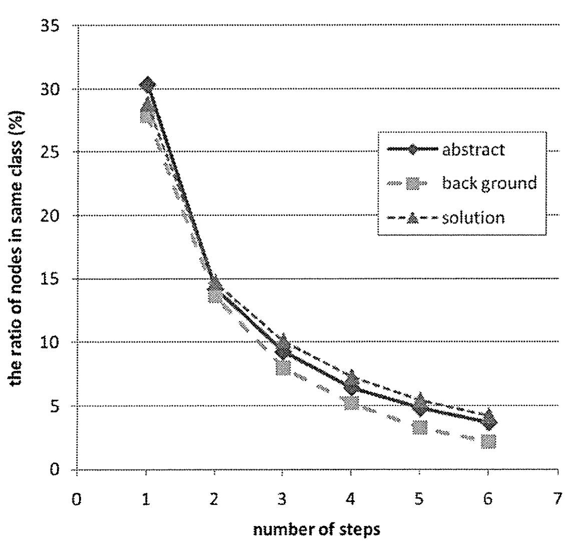

Figure 3. the ratio of nodes in same class

7. CONCLUSION

In this research, we evaluated the degree of similarities between incident documents obtained bottom-up and the links between existing classes granted top-down. In doing so, we made it possible to evaluate overall similarities regarding incident documents by using the method of network analysis. Moreover, it became clear that the use of the Cos coefficient or the Jaccard coefficient is appropriate for determining similarities in creating networks.

With regard to the background, the results of the analysis demonstrated that, compared with abstract and solution, existing classes are inadequate for representing the characteristics of documents and that there is a need to improve classes (figure 4).

By using the methods employed in this research, suggestions otherwise unobtainable through conventional methods can be made regarding the investigation of how classes should be.

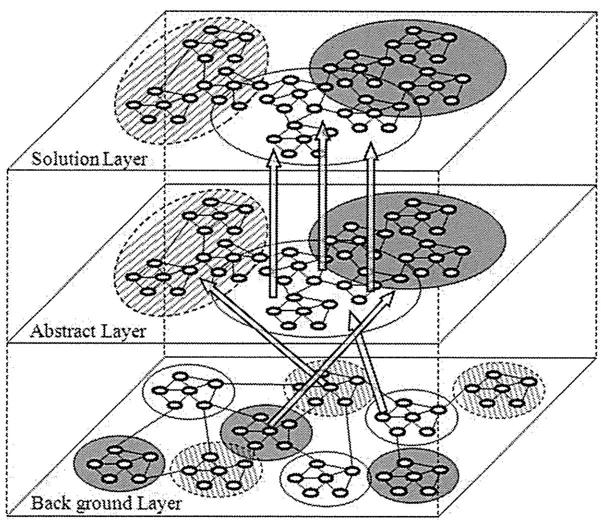

Figure 4. Structure of network in each layer

8. REFERENCES

Gandhi, T.K. and T. H. Lee (2010). Patient safety beyond the hospital. *The New England Journal of Medicine,* vol.363, pp.1000-1003.

Heinrich, H W. (1931). *Industrial accident prevention: A safety management approach.* McGraw-Hill Customer Service.

Huckvale, C., J. Car, M. Akiyama, S. Jaafar, T. Khoja, B. Farlow, S. Yaron, G. Locke and S. Whittaker (2010). Information technology for patient safety. *Quality & Safety in Health Care*, vol.19(2), pp. i25-i33.

Kajikawa, Y., J. Ohno, Y. Takeda, K. Matsushima, and H. Komiyama (2007). Creating an academic landscape of sustainability science: an analysis of the citation network. *Sustainability Science,* vol.2, pp.221–231.

Kaushal, R. and DW. (2002). Bates, Information technology and medication safety: what is the benefit ? *Quality & Safety in Health Care,* vol.11, pp.261-265.

Majeed, A., J. Car and A. Sheikh (2008). Accuracy and completeness of electronic patient records in primary care, *Fam Pract,* vol.25, pp.213-214.

Manning, C. D., and H. Schutze (2002). *Foundations of statistical natural language processing.* The MIT Press, London.

Matsuo, Y. and M. Ishizuka (2002). Keyword extraction from a document using word co-occurrence statistical information. *JSAI,* vol.17(3), pp.213-227.

Mcloughlin, V., J. Millar, S. Mattke, M. Franca, P.M. Jonsson, D. Somekh and D. Bates (2006). Selecting indicators for patient safety at the health system level in OECD countries. *International Journal for quality in Healthcare*, Sept 2006, pp.14-20.

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

Rasmussen, E. (1992). *Clustering algorithms, information retrieval: data structures and algorithms.* William B. Frakes and Ricardo Baeza-Yates.

Salton, G. (1983). *Introduction to Modern Information Retrieval.* Mcgraw-Hill College.

Shibata, N., Y. Kajikawa and I. Sakata (2011). Measuring relatedness between communities in citation network. *Journal of the American Society for Information Science and Technology*, in press

Shibata, N., Y. Kajikawa and I. Sakata (2010). Extracting the commercialization gap between science and technology. *Technological Forecasting and Social Change,* vol.77(7), pp. 1147-1155.

Uchida, M., N. Shibata, Y. Kajikawa, Y. Takeda, S. Shirayama and K. Matsushima (2009). Identifying the large-scale structure of the blogosphere. *Advances in Complex Systems,* vol.12, pp. 207-219.

http://mecab.sourceforge.net/

http://www.med-safe.jp/contents/english/index.html

http://www.who.int/patientsafety/implementation/taxonomy/en/

http://www.who.int/patientsafety/implementation/taxonomy/icps_statement_of_purpose.pdf

International Association for Management of Technology
IAMOT 2011 Proceedings

Information Science Linkage of Service Innovation

ICHIRO SAKATA

Policy Alternatives Research Institutes, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan,

isakata@ipr-ctr.t.u-tokyo.ac.jp

JUNICHIRO MORI

Innovation Policy Research Center, The University of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo, 113-8656, Japan

jmori@ipr-ctr.t.u-tokyo.ac.jp

NAOKI SHIBATA

Innovation Policy Research Center, The University of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo, 113-8656, Japan

shibata@ipr-ctr.t.u-tokyo.ac.jp

MASANORI AKIYAMA

Policy Alternatives Research Institutes, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan,

makiyama@pp.u-tokyo.ac.jp

YUYA KAJIKAWA

Innovation Policy Research Center, The University of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo, 113-8656, Japan

kajikawa@ipr-ctr.t.u-tokyo.ac.jp

Abstract

It is with this recognition that policies for strengthening international competitiveness regarding service innovation are being adopted by many countries. While planning and implementing these policies, what is required in essence is an objective analysis regarding the current status of knowledge related to this field and the linkage between science and innovation. However, the knowledge infrastructure of this kind is inadequate. The aim of this paper is to develop the way to identify the meta structure of knowledge and measure "information science linkage of service innovation". We use bibliometrics, co-citation network analysis and co-occurrence analysis to objectively identify them.

In the field of service innovation, our results show that there are mainly two groups of elements related to service innovation: applications of service innovation and basic theories for service innovation. In the field of service science, we also identified major knowledge groups such as machine learning, and information retrieval. Then we calculated the co-occurrence of characteristic terms in journal paper abstracts belonging to the SSME sub-clusters and author keywords for papers belonging to the sub-sub-clusters related to information science. We clarified which information sciences are used heavily in service innovation. We also determined which areas of innovation make heavy use of information science and which do not. In the field of medical care, the high value we obtained demonstrates that digital health and EHR research is being vigorously conducted worldwide. By contrast, in other fields related to medical treatment, such as mental healthcare and patient satisfaction, we found a big room to promote the spread of information science.

Overall, we have demonstrated the possibility of using bibliometrics to objectively identify the meta structure of knowledge and measure the semantic relationships between science and technology.

Keywords: SSME, Information Science, Science Linkage, Technology Roadmap

Introduction

It is widely recognized that the concept of service innovation is significant for innovation strategy and economic growth. It is with this recognition that policies for strengthening international competitiveness regarding service innovation are being adopted by many countries. While planning and implementing these policies, what is required in essence is an objective analysis regarding the current status of knowledge related to this field and the technological linkage between science and innovation. However, knowledge infrastructure of this kind is inadequate. The sense of concept SSME is so broad that there is not the common understanding about what is service innovation even among experts. Another reason is the lack of suitable indicator or the way of measurement. Even though, the indicators of science linkage which uses citation data of patents or corporate studies have been developed (Motohashi and Yun, 2007, Bonaccorsi and Thoma, 2007 as examples), they provide us limited information of semantic relationships of technologies. Therefore, using bibliometric methods, we shed light on the structure of knowledge pertaining to service innovation and the linkage or semantic similarity between service science and innovation.

The first aim of this paper is to create a knowledge landscape of service innovation and service science from a number of academic publications. In this paper, of the many definitions of service innovation, we focus on the concept "Service Science, Management, and Engineering (SSME)" proposed by IBM. This paper uses same data with our previous study of SSME (Shibata et al., 2009). As an indicator of service science, we use the data in the fields of information science, library science, and computer science, which are the most important basic knowledge for service innovation. We collect data of academic papers, create citation networks regarding papers as nodes and citations as links, categorize papers into sub-clusters or sub-sub-clusters, extract topics of each sub-cluster, and finally discuss the results with experts. The second aim is to calculate the degree of distance in the semantic connections between service innovation sub-clusters on the one hand and information science, library science, and computer science sub-clusters on the other. The degree of distance in these semantic connections may be employed as an indicator of the depth of the relationship between innovation and individual scientific technologies. This is referred to as "Information Science Linkage of Service Innovation." In addition, the relationship between information science and service innovation as viewed from the perspective of information science is referred to as "forward linkage," and the same relationship as viewed from the perspective of service innovation is defined as "backward linkage." In concrete terms, based on the likelihood that author keywords for journal papers within the clusters of information science, library science, and computer science will appear in abstracts of papers in the clusters within service innovation, we calculate the depth of the relationships between the various clusters that belong to both fields. Finally, we discuss the appropriate policy for promoting service innovation based on the knowledge infrastructure.

Methodology

First of all, the methodology for creating academic landscape is shown. Analyzing schema is depicted in Fig. 1. The step (1) is to collect the data of the knowledge domain. We collect citation data from the Science Citation Index Expanded (SCI-EXPANDED), the Social Sciences Citation Index (SSCI), and the Arts & Humanities Citation Index (A&HCI) compiled by the Institute for Scientific Information (ISI), which maintains citation databases covering thousands of academic journals and offers bibliographic database services, because these are three of the best sources for citation data. The problem, how we should define a research domain, is difficult to solve. One solution is to use a keyword that seems to represent the research domain. When we collect papers retrieved by the keyword, we can make the corpus for the research domain.

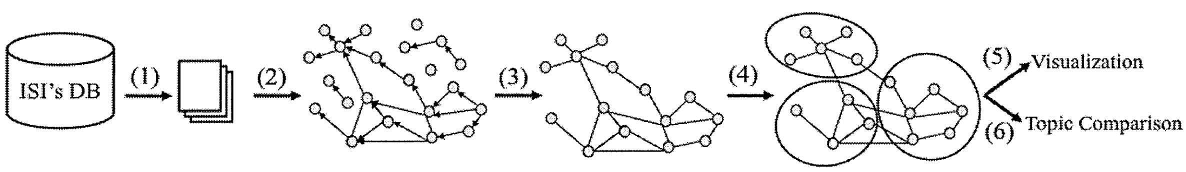

Figure 1. Methodology proposed in this paper.

The step (2) is to make citation networks for each year. We construct citation networks by regarding papers as nodes and intercitations as links. The network created for each year facilitates a chronological analysis of citation networks. According to a previous study, intercitation, which is also sometimes known as direct-citation, is the best way to detect emerging trends (Shibata et al., 2009, Shibata et al., 2010, Sakata et al., 2010). In network analysis, only the data of the largest component on the graph was used, because our study focuses on the relationships among documents, and we therefore want to eliminate from our study those not linked with any others in step (3).

After extracting the largest connected component, in step (4), the network is divided into clusters using the topological clustering method (Newman, 2004), which does not need the number of clusters by users. Newman's algorithm discovers tightly knit clusters with a high density of links within cluster. In step (5, 6), experts in the research domain assign a name to each cluster manually after they had seen titles and abstracts of the papers in each cluster, supported by the methodology of visualization developed by Adai et al. (2004).

Second, we calculate "Information Science Linkage of Service Innovation." The linkage is defined as the degree of distance in the semantic connections between service innovation clusters on the one hand and information science clusters on the other. The degree of distance in these semantic connections may be employed as an indicator of the depth of the relationship among different research fields.

For respective clusters of information science, we calculate the semantic connection to each cluster of service innovation. We define this linkage, as viewed from the perspective of information science to service innovation, as "forward linkage". And the inverse linkage, as viewed from the perspective of service innovation to information science, is called as "backward linkage".

More formally, let *linkage(I,S)* be the linkage between a cluster I of information science cluster and a cluster S of service innovation:

$$linkage(I,S)=\begin{cases}\sum_{W}\dfrac{freq_s(w_I)}{|A_s|} & forward\ linkage\\[2ex] \sum_{W}\dfrac{freq_s(w_I)}{|W_I|} & backward\ linkage\end{cases},$$

where W_I is a set of author keywords, which are assigned to a paper by its author(s), of papers in I, w_I is an individual keyword in W_I, A_s is a set of abstracts of papers in S, and $freq_s(w_I)$ denotes a term frequency of w_I in A_s. The forward linkage is normalized with the size of A_s so that linkages from one of information science clusters to service science clusters are uniformly evaluated. And the backward linkage is normalized with the size of W_I.

The number of abstracts in each cluster of service science is shown in Table 1 and the number of author keywords in each cluster of information science is shown in Table 2.

In our experiment, we exclude an author keyword that its frequency is less than a certain threshold. We set the threshold as 100 based on our preliminary experiment.

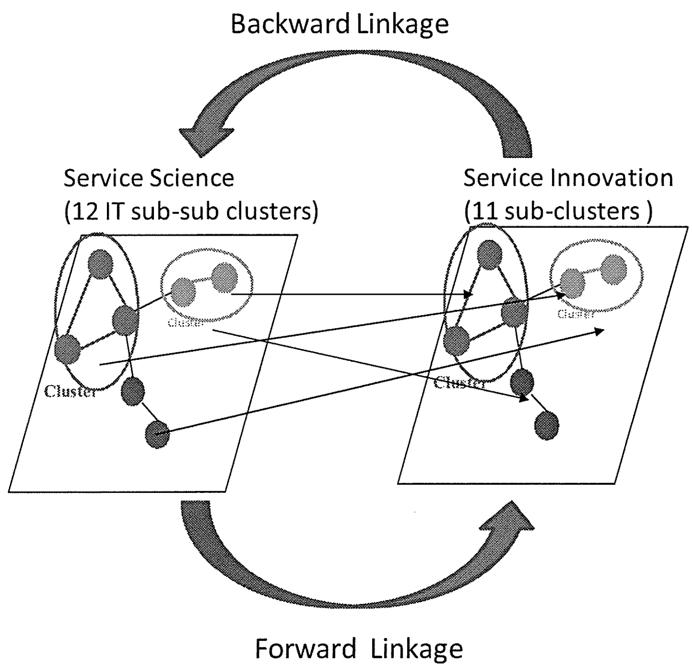

Figure 2.Concept of information science linkage

Results

With respect to service innovation, the number of academic papers that form the target of this analysis stands at 54,928. Our results shows that there are mainly two groups of elements related to service innovation: applications of service innovation such as health and medical care, IT, and public service, and basic theories for service innovation such as management, ecosystem, and QoS (Table 1).

Table1.Major sub-clusters of service innovation (SSME)

Service Cluster	Label	# of Abstracts
1	Management	1783
2	Medical Care	1660
3	Mental Health Care	1276
4	Ecosystem	901
5	Quality of Service (QoS)	903
6	Public Service	837
7	Public Medical Care	627
8	IT and Web	454
9	Patient Satisfaction	351
10	Clinical Pharmacy	305
11	Telemedicine	308

The number of academic papers related to information science, library science, and computer science stands at 314,806. Major sub-clusters include (1) machine learning, neural network, computer vision and computer graphics, (2) artificial intelligence, network, information retrieval, information theory and database, (3)distributed or parallel computing, computer architecture and information system, (4)fuzzy, (5)bioinformatics, (6)security and cryptography, (7)library and information science, (8)computer physics, (9)math application, (10)telecommunication, (11) reliability, (12)graphics, display and color, (13)computer and geo science, (14)health information, (15)circuit device. Our analysis targets the top three sub-clusters which include more than 60,000 papers. Major sub-sub-clusters of the top three sub-clusters include machine learning, computer vision, neuroinformatics, computer graphics, artificial intelligence, telecommunication network, information retrieval, information theory, distributed computing, computer architecture, information system, and multimedia (Table 2).

Table2.Major sub-sub clusters of service science

Information Science Cluster	**Label**	**# of Keywords**
1	Machine Learning	825
2	Computer Vision	657
3	Neuroinformatics	782
4	Computer Graphics	552
5	Artificial Intelligence	722
6	Telecommunication	615
7	Information Retrieval	571
8	Information Theory	513
9	Distributed Computing	576
10	Computer Architecture	631
11	Information System	550
12	Multimedia	592

We calculate the co-occurrence of characteristic terms in paper abstracts belonging to the 11 SSME sub-clusters and author keywords for essays belonging to the 12 sub-sub-clusters belonging to information science. The co-occurrence matrix shows the raw data of co-occurrence (Table3). The biggest number is 12,816 and the smallest number is 1,014. The pair which has biggest number is the pair of innovation sub-cluster #1 (Management) and information science sub-cluster #5 (artificial intelligence).

Table 3.Co-occurrence matrix

		Service Innovation Cluster										
		1	2	3	4	5	6	7	8	9	10	11
Info.	1	11,716	8,726	5,137	5,051	11,094	4,213	3,566	3,951	1,870	1,613	1,923
Science	2	8,773	6,618	3,920	3,765	7,963	3,184	2,748	2,816	1,392	1,207	1,488
Cluster	3	10,925	8,407	4,866	4,972	10,577	4,032	3,426	3,798	1,827	1,607	1,943
	4	7,011	5,398	2,944	3,104	7,740	2,550	2,268	2,353	1,090	1,014	1,111
	5	12,816	9,814	6,083	5,311	11,933	4,631	3,935	4,685	2,108	1,849	2,225
	6	7,779	6,012	3,449	3,587	11,038	3,003	2,483	2,740	1,298	1,171	1,358
	7	12,127	9,172	5,683	4,765	10,529	4,471	3,633	4,320	1,955	1,700	2,078
	8	7,493	5,927	3,341	3,293	10,101	2,990	2,573	2,477	1,263	1,127	1,357
	9	10,048	8,444	5,107	4,369	12,177	3,798	3,367	3,654	1,702	1,584	1,731
	10	8,823	7,078	4,029	3,641	11,353	3,227	2,818	3,139	1,422	1,388	1,543
	11	12,778	9,694	6,015	5,186	11,588	4,672	3,945	4,555	2,081	1,833	2,210
	12	10,775	7,390	4,742	4,448	11,456	3,940	3,140	4,039	1,666	1,380	1,821

Then we normalize the raw data, using the size of abstracts or author keywords as denominator. It became clear that artificial intelligence, information retrieval, and distributed computing values were high for forward linkage (figure2). It may be considered that these sciences are being applied heavily in cutting-edge service innovation. Values in the three fields of quality of service (QoS), management, and medical care were high for backward linkage (figure3). It may be considered that information science is being used very heavily in the context of innovations in these fields. On the other hand, values in the three fields of public service, mental healthcare, and patient satisfaction were low. We believe that the degree to which information science is being applied is lower in these fields than in others.

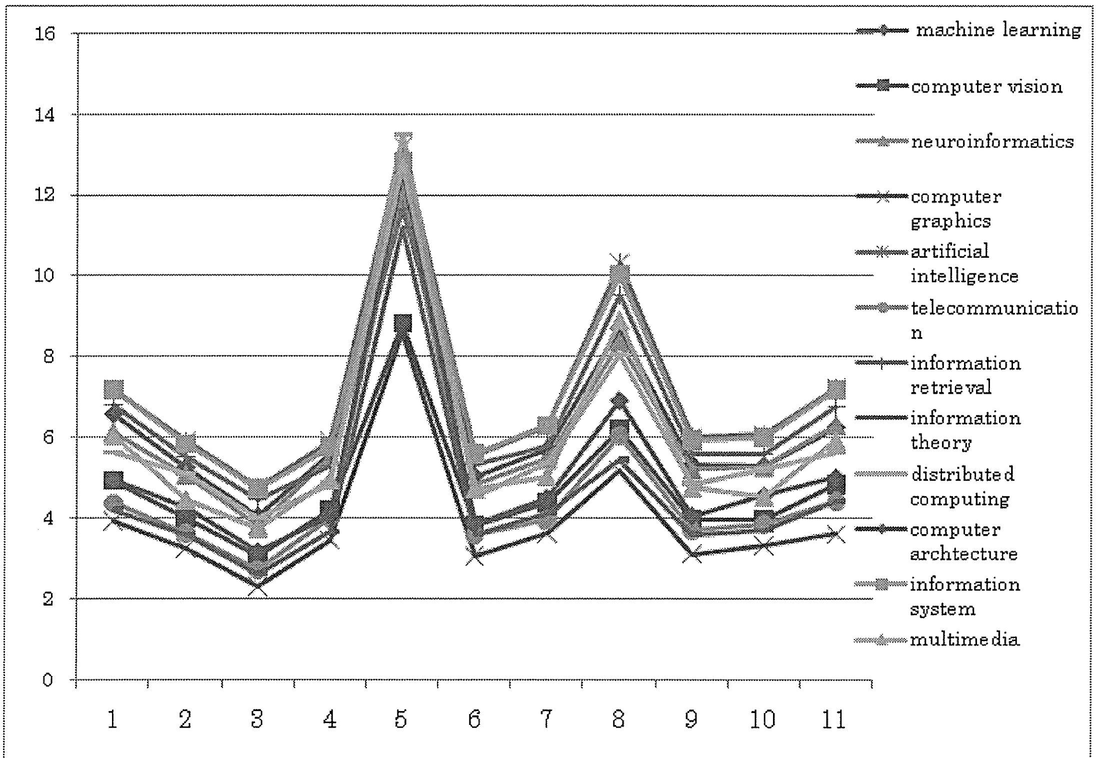

Figure 2. Forward Linkage

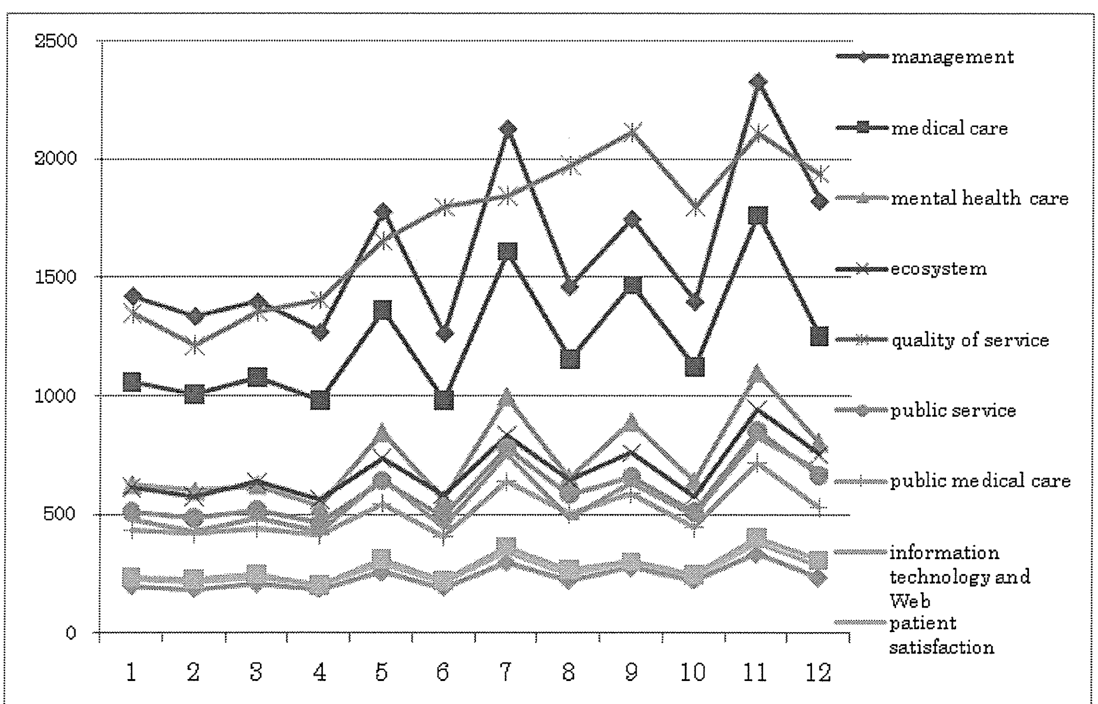

Figure 3. Backward Linkage

Discussion

First, the results of the analysis described above clarified which information sciences are used heavily in service innovation. It is known that these fields possess deep links with forms of information engineering such as data mining, the web, and cloud computing. Moreover, these fields are similar to the list of technologies specified as particularly important by experts in the Technology Strategy Roadmap of service engineering issued by the Japanese organization NEDO (NEDO, 2008). The roadmap mentioned forty nine technology elements and the relationship between technology and industry, even though these descriptions lack concreteness of linkage. The technologies which are indentified in both NEDO roadmap and our results include machine learning, neural network, distributed computing, information theory and multimedia. Our method, which used bibliometrics, produced results similar to those obtained by multiple experts working over a long period of time. In a rapidly changing field like information science, it is important to regularly obtain updated information; however, this is difficult using a method, typically referred to as the T-plan, in which experts form a consensus. Our method has the potential to contribute to technology strategies in a field in which technology continues to progress rapidly.

Second, we determined which areas of innovation make heavy use of information science and which do not. In the field of medical care, the high value we obtained demonstrates that digital health and EHR research is being vigorously conducted worldwide (Oren et al., 2003, Eslami et al., 2007, Huckvale et al., 2010). By contrast, in other fields related to medical treatment, such as mental healthcare and patient satisfaction, usage of information science is relatively low. It is believed that information science linkage improvement strategies have had a significant effect in these related fields. Public service—a field for which we judged use to be low—has been specified in OECD's Innovation Strategy (2010) as one toward which it is particularly important that public sector takes innovation-conscious attitudes. Our analysis suggests that it is important to develop strategies for accelerating the spread of information science in public sector.

Conclusion

It is with this recognition that policies for strengthening international competitiveness regarding service innovation are being adopted by many countries. While planning and implementing these policies, what is required in essence is an objective analysis regarding the current status of knowledge related to this field and the linkage between science and innovation. However, the knowledge infrastructure of this kind is inadequate. Therefore, we developed the way to identify the meta structure of knowledge and measure “information science linkage of service innovation”.

With respect to service innovation, our results show that there are mainly two groups of elements related to service innovation: applications of service innovation and basic theories for service innovation. In the field of service science, we also identified major knowledge groups such as machine learning, pattern recognition, computer vision, objectively. Then we calculated the co-occurrence of characteristic terms in paper abstracts belonging to the SSME sub-clusters and author keywords for papers belonging to the sub-sub-clusters related to information science. We clarified which information sciences are used heavily in service innovation. It is known that these fields possess deep links with forms of information engineering such as data mining, the web, and cloud computing. We also determined which areas of innovation make heavy use of information science and which do not. In the field of medical care, the high value we obtained demonstrates that digital health and EHR research is being vigorously conducted worldwide. By contrast, in other fields related to medical treatment, such as mental healthcare and patient satisfaction, we found a big room to promote the spread of information science in public sector.

Overall, we have demonstrated the possibility of using bibliometrics to objectively identify the meta structure of knowledge and measure semantic relationships between science and technology.

Acknowledgement

This paper was partly supported by a grant from the Economic and Social Research Institute (ESRI), Cabinet Office, Government of Japan. The Views contained here reflect those of authors.

References

Adai, A.T., 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*, vol. 340(1), pp.179-190, 2004.

Bonaccorsi, A. And G. Thoma, "Institutional complementarity and inventive performance in nano science and technology," *Research policy* 36(6), pp.813-831, 2007.

Eslami, S., A. Abu-Hanna and NF. de Keizer, "Evaluation of outpatient computerrized physian medication order entry sysetm: a systematic review, " J Am Med Inform Assoc. 14, pp.400-406, 2007.

Huckvale, C., J. Car, M. Akiyama, S. Jaafar, T. Khoja, A. Binkhalid, A. Sheikh and A. Majeed, "Information technology for patient safety," *Qual Saf Health Care* 19, pp.i25-i33, 2010.

Motohashi, K. And X. Yun, "China's innovation system reform and growing industry and science linkages," *Research Policy* 36(8), pp.1251-1260, 2007.

New Energy and Industrial Technology Development Organization (NEDO), "NEDO Technology Strategy Roadmap" 2008, http://www.nedo.go.jp/roadmap/2008/index.html

Newman, M. E. J., "Fast algorithm for detecting community structure in networks," *Physical Review E*, 69, 066133, 2004.

OECD, *The OECD Innovation Strategy: getting a Head Start on Tommorow*, OECD press, 2010.

Oren E., ER. Shaffer, BJ. Guigliemo, "Impact of emerging technologies on medication errors and averse drug events, *Am J Health Syst Pham* 60,pp.1447-1458, 2003.

Sakata, I., H.Sasaki, Y.Kajikawa, K.Fujisue and A.Morita, "Systematic identification of academic knowledge in patent & innovation research," *proc. of IEEE International Conference on Management Science and Information Engineering 2010* (ICMSIE2010) in ZhengZhou, China (December 17-19,2010)

Shibata, N., Y. Kajikawa, Y. Takeda and K. Matsushima, "Comparative study on methods of detecting research fronts using different types of citation," *Journal of the American Society for Information Science and Technology*, vol.60(3),pp. 571-580, 2009.