Human behavior analysis in the production and consumption of scientific knowledge across regions

A case study on publications in Scopus

Received 17 April 2017
Accepted 2 June 2017

Muhammad Awais Qasim
Department of Computer Science, Information Technology University, Lahore, Pakistan
Saeed Ul Hassan
Information Technology University, Lahore, Pakistan
Naif Radi Aljohani
Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, and
Miltiadis D. Lytras
The American College of Greece, Athens, Greece

Abstract

Purpose – The latest developments in Data Science and in advanced Scientometrics set a very challenging context for the analysis and the understanding of human behavior toward the design of value adding library services and sophisticated information systems. The purpose of this paper is to present an innovative research that integrates the creation and the consumption of scientific knowledge across regions. From a human behavior point of view, this is significant since it provides an advanced decision-making layer for bringing together researchers from all over the world.

Design/methodology/approach – More specific in this paper, the authors analyze the production and consumption of scientific knowledge across the regions in an important field of sustainable and renewable energy – using publications and citations data indexed in Scopus. As a case study, the authors select the USA a major producer of scientific publications in the field. At first, the authors identify the topics produced by the USA. Further topics produced by the scientific communities outside the USA that consume the knowledge produced by the USA are identified. The authors generate topics by employing the proposed topic model with distance matrix – an extension of classic latent Dirichlet allocation model.

Findings – The results show that research topics produced by the USA are consumed in different international contexts, interestingly. Consuming the knowledge produced by the USA, Chinese scientific community heavily produces topics related to biomass – to produce renewable energy. In contrast, Japanese scientific community produces topics related to fuel cell – used for the production of hybrid and electronic vehicles. Whereas the Taiwanese scientific community shows remarkable competency in solar cells. Among the European nations, while the German scientific community produces topics related to photovoltaic, the French scientific community covers topics related to Energy Storage and Green Chemistry. The authors believe that such analyses may be helpful in establishing more effective multi-national research collaborations by understating the actual consumption of produced knowledge.

Practical implications – Overall, the study provides a new dimension to comprehensively understand production and consumption of knowledge using scientific literature. From a human behavior analysis view in the context of sophisticated library systems, this is a significant contribution.

Originality/value – The use of advanced Data Mining and computing methods for deriving critical insights for the use of scientific knowledge is a bold action toward the global knowledge society vision.

Keywords Library services, Data Mining, Social networks, Scientometrics, Data Science, Human behaviour analysis

Paper type Research paper

Library Hi Tech
Vol. 35 No. 4, 2017
pp. 577-587

0737-8831
DOI 10.1108/LHT-04-2017-0076

1. Human behavior analysis for scientific exchange and knowledge transfer

The design of sophisticated library services in our days is a strategic objective for a new era of academic excellence and scientific performance. The variety of enabling technologies, including data Mining, Learning Analytics, Cloud Computing, Recommender Systems, Data Science, Internet of Things, revolutionize the value adding propositions. A key question that should be answered is how can we vision the next-generation library systems enabled by Hi Tech. To our understanding, this special issue is about analyzing design principles for library services based on extensive analysis and integration of human behavior. For this purpose, we set in the focus of our approach a significant context: the knowledge flows across regions, related to scientific publications. It is not only about mining and applying advanced Data Mining and data analytics methods, but also including clustering, classification, association rules, link analysis, text mining and sentiment analysis in the practice of modern libraries. Most of all, it is about integrating human behavior in transparent services that foster the international collaboration and permit the dissemination of scientific know-how for humanitarian visions. The development of a Global Knowledge Society is in the focus of our research. We promote the role of libraries as key disseminators, and we analyze the significance of Data Science for the detailed analysis of requirements and use case scenarios that can foster the international collaboration of researchers based on common understanding and joint contributions in scientific domains.

In the last few decades, scientists have made a remarkable progress in the production of knowledge resources by interacting with other scientists across the globe – in order to develop new ideas and technologies (Jones *et al.*, 2008; Hassan and Haddawy, 2015a). As a result of these interactions, the generation and retention of knowledge is not solely dependent on a single isolated entity, but rather depends upon the knowledge flow between different dynamic entities (Yan *et al.*, 2012). A significant among of research has been done to quantitatively measure knowledge production and consumption – using absolute citation counts as a measurement tool. No doubt, these researches have made a sound base to study the interdisciplinary, inter-organizational and international knowledge flows, but have not addressed utmost important need of semantic understanding of production and consumption of scientific knowledge.

The main objective of this paper is to semantically analyze the production and consumption of scientific knowledge across the countries – in a very important field of "Sustainable and Renewable Energy," which has recently become an immense business in many parts of the world due to climate change concerns and increase in global carbon emission (Zahari and Esa, 2016). Alone, in 2015, global investment in renewable and sustainable energy raised to $285.9 billion[1] – taking it above the previous record of $278.5 billion reached in 2011. Recently, China has increased research and development (R&D) spending on renewable energy to 4 percent by investing $2.5 billion, making China the largest producer of solar cells in the world. Similarly, Germany is producing its one-third of the total energy from the renewable resources. Last but not the least, German Government has allocated approximately €3.5 billion funding for the R&D in energy sector during 2011-2014. Given the crucial need for the economies to move toward sustainable renewable energy resources and to learn from each other, it also becomes vital to analyze that what research themes are produced within the field of sustainable and renewable energy and how they are consumed or transferred across the geographical and disciplinary boundaries.

To our understanding, modern libraries are key enablers of advanced knowledge dissemination. They can serve as hubs for analyzing significant aspects of human behavior and the emerging technologies with their advanced capabilities for profiling, clustering and mining that allow significant knowledge transformations. The development of a knowledge ecosystem for the knowledge integration and knowledge dissemination across regions is a bold initiative. Our research exactly sets this problem in the focus. The inquiry to understand

how scientific knowledge is produced and consumed can provide useful insights for the exploitation of human behavior in the context of library and information sciences. Libraries should be seen not only as facilitators of international collaboration, but also as means for content, knowledge delivery.

In this paper, we select the USA as main scientific knowledge producer in the field of sustainable and renewable energy. We cluster the scientific publications produced by the USA into thematic areas, using keywords from paper titles, author-defined keywords and abstract. Further, we analyze (by clustering) the publications produced by the countries other than the USA that cite original publications produced by the USA. This helps to analyze the consumption of knowledge produced by the USA by other countries. The rest of the paper is organized as follows: Section 2 presents related literature review on scientific knowledge flows originated by publication and citation data. Section 3 presents the employed methodology. Section 4 presents a case study to analyze the production and consumption of scientific knowledge in the field of sustainable and renewable energy. Finally, Section 5 presents concluding remarks along with the direction of future work.

2. Literature review

Significant amount of research work has been done to analyze knowledge flows by using citation and bibliometric data. Porter and Chubin (1985) were among the first to study interdisciplinary knowledge flows with the help of bibliometric data. They used citations as an indicator of knowledge flow between different fields in scientific literature. According to Borgman (1990), citation links in a citation network could be used to find the impact of research in a particular area. He described that a journal having the larger number of citations has greater impact as compared to the journal having smaller number of citations. Trajtenberg (1990) argued that patent citations reflect the quality of innovation in science and technology. He also established relationship between patent citations and R&D expenditure. Jaffe *et al.* (1993) used geographical location data set of patent citations alongside the cited patents to find knowledge flow spillovers geographically. By using existing approaches, they were unable to find patents that were highly cited and had large impact on technology and economy. In general, it is argued that citations are good indicator of knowledge flow in scientific and research environments. Also, citation links among institutions, journals and authors denote knowledge flow from cited entity to citing entity (Stigler, 1994; Lockett and McWilliams, 2005; Yan *et al.*, 2013; Guevara *et al.*, 2016; Khasseh *et al.*, 2017).

Ingwersen *et al.* (2000) utilized percent international citations as an indicator of knowledge flow between the institutions. The percent international citations are the proportion of citation that an institute receives from the institutes outside the country with the total citation it receives. Cohen *et al.* (2002) discovered that knowledge is transferred through research interactions, meetings in conferences, personal contacts and citations in publications. They found through a survey conducted among R&D managers of the US firms that research articles or publications are the dominant source of knowledge flow with 41 percent of respondents choosing this option. Rinia *et al.* (2002) used bibliometric data to find interdisciplinary research impact. They found that patent citations alongside with bibliometric indicators can represent useful information between different fields of science.

Alcacer and Gittelman (2006) used citations data to find knowledge diffusion. They analyzed citations with respect to geographical location, scientific innovation and self-citation of examiner and inventor of citations. Zhuge (2006) described that in scientific environments, research interactions and development are heavily dependent upon understanding of knowledge flows between scientists. He explored different knowledge flow networks and presented a methodology to discover knowledge flow in scientific literature. Nicholas Joint (2008) examined the impact of library and information science issues on higher education policy related with

quality research in the UK using citation analysis. Nomaler and Verspagen (2008) showed that a number of citations in a scientific literature could indicate the process of knowledge flows. They proposed a measure to find interdisciplinary knowledge flow and its diffusion into economic system, but their analysis did not define relation between specific nature of knowledge and domain in which knowledge is applied. Zhou (2010) found the impact of the research community within a specific discipline by analyzing citation exchange among the journals. Their study showed that Chinese-based journals are not specialized as compared to their international counterparts in the disciplines of science, economics and politics. Yuan and Hua (2011) measured scholarly impact of library and information science open access journals on the basis of citations. They found open access journals are substantial part of scholarly communication. Tsay and Shu (2011) found the relationship between different social sciences subjects using citation database. Lochan Jena *et al.* (2012) used bibliometric data set of electronic library to study scholarly communication and found contributor's circulation is restricted to limited nations. Also, Tsay (2013) explored the knowledge flow between information science and other subjects using bibliometric and citation database. Hassan and Haddawy (2013) introduced a quantitative approach to find scholarly impact between different countries. A new quantitative technique was defined to find out the international scholarly impact of the countries by using bibliometric and citation data in the field of energy.

Zhang (2013) explored different methodologies to find scientific trends of production and consumption of scientific knowledge flow in the area of physics. He proposed knowledge diffusion proxy algorithm to find out the knowledge flow from producers to consumers. Nagaoka and Yamauchi (2015) used citations as an index to find knowledge flow in literature. They found the patent citations to non-patent literature are noisy. More recently – on semantic understanding of knowledge flows, Hassan and Haddawy (2015b) proposed a methodology to analyze knowledge flow across countries. Using an extended version of latent Dirichlet allocation model, the authors identify the research topics produced by a given country in a given research area and present a semantic understanding of research topics in the cited topics across the countries. More recently, Ding *et al.* (2017) explored paper characteristics that facilitate the flow of knowledge from science to technology by exploiting patent-to-citation data.

3. Methodology

We analyze the production and consumption of knowledge across the countries in the field of sustainable and renewable energy. At first, we identify papers produced by the source country in the field. We cluster these papers into relevant groups to analyze different themes produced by the source country – we name them production topics. Further, we identify papers that cite these production topics. Next, we cluster citing papers to analyze the consumption of knowledge into relevant themes – we name them consumption topics. We employ our proposed TDM model (Hassan and Haddawy, 2015b), an extension of the classic LDA model (Blei *et al.*, 2003), to generate production and consumption topics from the publication data set. The classic LDA model is based upon the idea that probability distribution of words in a paper can formulate different topics. At first, using Dirichlet distribution, a distribution over topics is sampled for each paper. Further, topics are associated with each word in the paper based on the probabilistic distribution. Finally, words are sampled using multinomial spread over words particular to a given topic.

The LDA provides topic vectors, where each value in the vector corresponds to a word in the topic under consideration. For example, a given vector t_i (w_1: 0.2, w_2: 0.1, w_3: 0.8, …, w_n: 0.6) shows probability of each word w_i in topic t_i. Further, for each paper p_i, we generate a vector that shows probabilistic relation of each w_i in p_i for a given topic t_i: p_1 (w_1: 0.1, w_2: 0.0, w_3: 0.2, …, w_n: 0.3). Note that the probability of given word in p_i for t_i would be 0 if the word does not exist in paper p_i. Also, all unique words across the topics formulate vocabulary V. Further, we

use Minkowski distance between a given paper p_i and a topic t_i, to group papers that belong to a given topic (see the following equation):

$$D = \sqrt{\sum_{i=1}^{n} |a_i - t_i|^2}, \quad (1)$$

where a_i is the probability of word w_i in p_i for a given topic t_i and n the total number of words in vocabulary V. Further, to obtain relevant papers belonging to topic t_i, a threshold TH, with a percentage between the least and the most distance of p_i from t_i, is applied – our experiments show highest F-measure on a $TH = 25$ percent. The optimum numbers of topics are computed by minimizing inter-topic similarity and maximizing intra-topic similarity by using Jaccard similarity index (see the following equation), where S is set of topics and t_i and t_j are research topics, respectively:

$$InterTopicSim = \min\{(1 - dis(t_i, t_j)) | t_i, t_j \in S\}, \quad (2)$$

Further, to obtain intra-topic Jaccard similarity of topic t_i, we compute centroid d_t using Equation (3). Then, intra-topic similarity is measured by taking the average similarity between d_t and all elements in t_i, as shown in Equation (4):

$$\underset{t}{\text{argmax}} \quad \frac{\sum_{d \in t} sim\ (d_t, d_j)}{n} \quad | d_t, d_j \in t \quad (3)$$

and:

$$IntraTopicSim(t) = \frac{\sum_{j=1}^{n} sim(d_t, d_j)}{n} | d_t, d_j \in t, \quad (4)$$

where d_j is the jth element of a given topic; and n the total number of elements in a given topic.

4. Case study: production and consumption of scientific knowledge in the field of "sustainable and renewable energy"

4.1 Data set

Our data set consists of 9,090 papers which have at least one author affiliated from the USA in the field of sustainable and renewable energy, indexed in All Science Journal Classification used by Scopus, during 1996-2009 – also these publications are only cited by the papers which have at least one author affiliated from outside the USA. We refer this data set as the scientific knowledge produced by the USA in the field. Further, we select all 23,838 papers (authored by the researchers from outside the USA) that cite these 9,090 papers during the same time window. We refer the second data set as the consumer of the knowledge produced by the USA in the field.

The following steps are taken to process the data. In data cleaning phase, we remove all HTML tags from title, abstract and author-defined keywords of papers. Further, we perform tokenization, at the same time, we remove stop words, special characters, such as punctuation, question mark, etc., from the data. The Porter stemming is performed on the text (http://tartarus.org/martin/PorterStemmer/). Using "wordle" topics is visualized (www.wordle.net/). Finally, the data are fed to our TDM model for processing.

4.2 International consumption of knowledge produced by the USA in the field

Figure 1 shows the knowledge consumed by the world in the field of sustainable and renewable energy, produced by the USA during 1996-2009. The red color intensity in the

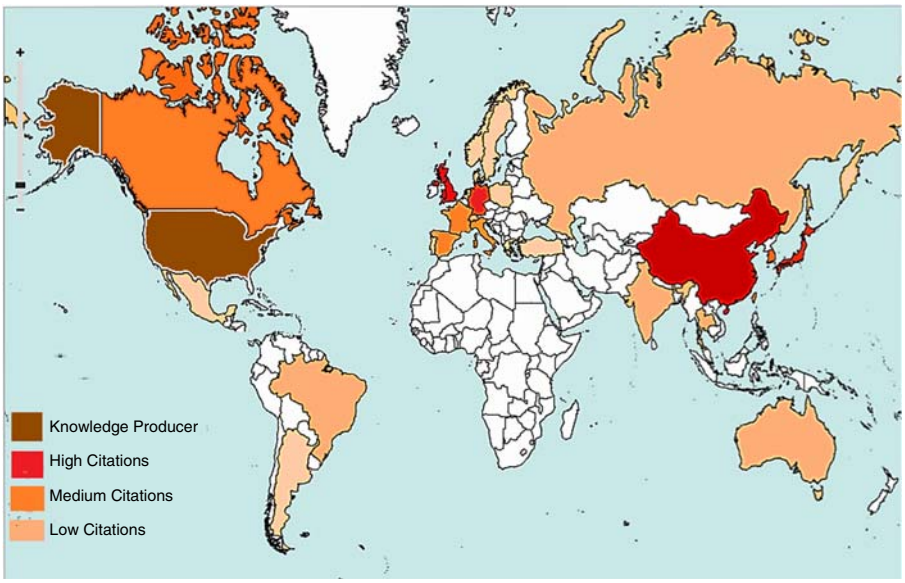

Figure 1.
International consumption of the knowledge in the field of sustainable and renewable energy produced by the USA during 1996-2009

map shows the proportion of citations made by each country to the scientific knowledge produced by the USA. Among the top ten nations, that consume the knowledge produced by the USA, 20 percent is consumed by China. Followed by China are the UK (16 percent), Germany (11 percent), Canada (10 percent), France (9 percent), Australia (8 percent), South Korea (8 percent), Taiwan (8 percent), Japan (7 percent) and Brazil (6 percent). Interestingly, we find many fast-growing economies in Asia that are listed among the top consumers of the US knowledge. Despite the relatively small size compared to many European nations, countries like Taiwan and South Korea are heavily consuming the US knowledge and produce new knowledge resources in the field. This clearly indicates that emerging economies heavily consume the knowledge produced by technically advanced knowledge economies.

4.3 Topics produced by the USA that are consumed by the international community

Figure 2 shows production topics of the USA in the selected field. A word cloud represents each topic, where the most frequent terms in each cloud is represented with a relative large font size. The cluster no. 1 appears to be the largest topic with 34 percent of total 9,090 papers that are internationally cited outside the USA. This shows that geothermal renewable energy topic is the largest knowledge resource consumed by the international community. The cluster no. 2 represents the themes related to fuel cells – solid oxide fuel cell and polymer fuel cells. The cluster no. 3 is about the storage and production of sustainable renewables. The cluster no. 4 is related to lithium battery production and power efficiencies. Finally, the cluster no. 5 represents themes related to the production and efficiency of thin film.

4.4 Topics produced in the field of sustainable and renewable energy after consuming the knowledge of cluster no. 1 in Figure 2

Further, we show the knowledge produced by the nations after consuming the knowledge produced by the USA in the field. To do so, we selected largest consumed topic i.e. cluster no. 1

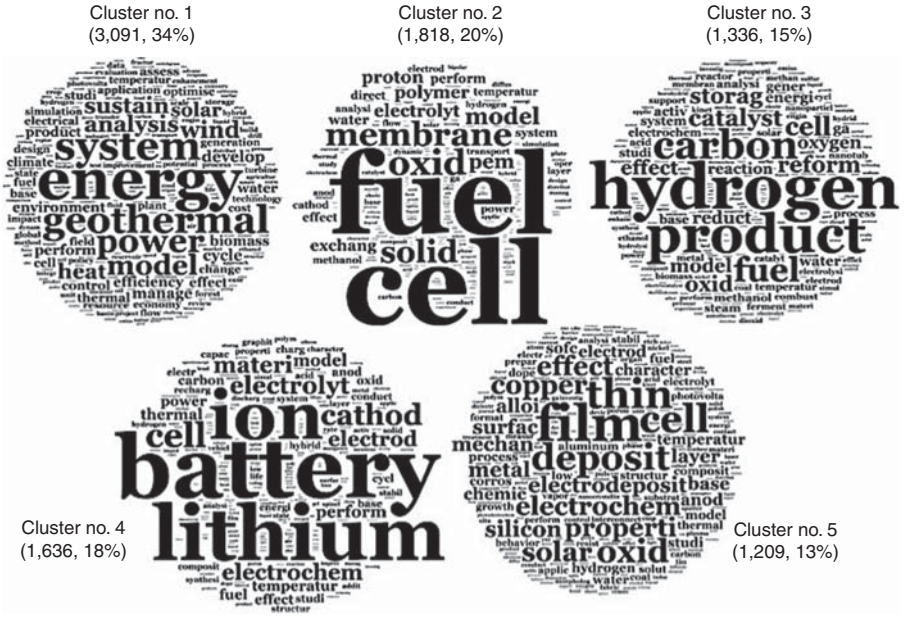

Figure 2.
Research topics of the scientific knowledge produced by the USA during 1996-2009

(see Figure 2) and fetch 23,838 papers that cited this topic. Figure 3 shows the topics that consume the knowledge produced by the USA. The cluster no. 1 (in Figure 3) shows research themes related to the area of nanotechnology in relation to boost the geothermal energy. Note nanotechnology is in use to cut geothermal energy cost and increase efficiency. The cluster no.

Figure 3.
Research topics produced by the regions after consuming the knowledge produced by the USA (in Figure 2 cluster no. 1) during 1996-2009

2 represents the research topics in the area of electrochemical. This cluster presents production of the materials by using chemical processes to increase the performance of renewable and sustainable energy resources. The cluster no. 3 represents the research topics related to the solar cell. China is the biggest producer of solar cell with 43.2 gigawatts of solar capacity (www.technologyreview.com/). The cluster no. 4 represents the topics related to fuel cells in relation to geothermal. Note that hybrid energy systems, that include fuel cells and geothermal resources, are developed to improve the efficiency by using low-grade heat from the fuel cells. The cluster no. 5 represents the research topics related to thin film. Finally, cluster no. 6 discusses the production of renewable energy by using biomass. Note that bioenergy provides 10 percent of the world's total global energy supply[2].

4.5 Research topics produced by major economies after consuming the knowledge produced by the USA in Figure 2 cluster no. 1

This section presents a detailed discussion on the production of the knowledge by major economies after consuming the knowledge produced by the USA in Figure 2 cluster no. 1. At first, we discuss the topics produced by the Chinese scientific community, follow by Japan and Taiwan. Later, we also present topics produced by Germany, France and Canada.

4.5.1 Research topics produced by China. The Chinese scientific community produces diverse research topics by consuming the knowledge related to the geothermal-related topics produced by the USA: photovoltaic energy and biomass are the main topics. The reason China being the largest consumer of knowledge produced in the field is due to its huge R&D budget allocated for the renewable and sustainable energy sector – resulting a large number of scientists working in the field. According to energy research knowledge center (https://setis.ec.europa.eu/energy-research/country/china), China is the world's second largest investor in R&D, next to the USA. In 2010, the total R&D expenditure has reached €70.6 billion and it is expected to reach €112 billion (2.5 percent of GDP) in 2020. China invested €205 billion in renewable energy projects in 2011 and plans to spend €367 billion on clean energy investments by 2015.

4.5.2 Research topics produced by Japan. In contrast to China, a large portion of Japanese research falls into the area of fuel cells. In last few decades, due to its automobile industry shift toward hybrid and electronic vehicles, Japan has done remarkable research in the field of fuel cells. Note that different kind of fuel cells like polymer electrolyte fuel cells, phosphoric acid fuel cells, and molten carbonate fuel cells are produced in Japan. The Japanese Government is heavily investing in fuel cells sector. It has allocated ¥142.3 billion in its 2005 budget for science and technology promotion, with fuel cells accounting for ¥35.4 billion, which is around 25 percent of the total budget. Japanese automobile industry plans to shift on fuel-cells-based vehicles by reducing 90 percent carbon emission before 2050[3].

4.5.3 Research topics produced by Taiwan. The Taiwanese solar cell manufacturing industry is second in the world, only next to China, with the revenue of US$5.75 billion in 2014. Research topics produced by Taiwan are related to photovoltaic cells that generate electric power by using solar cells to convert energy from the sun into a flow of electrons. Note that Taiwan's renewable energy sector has grown in value by 195 percent between 2008 and 2014, increasing to over US$15 billion in annual[4].

4.5.4 Research topics produced by Germany, France and Canada. The German scholarly community produces topics related to photovoltaic solar cells. Note that Germany is producing its one-third of total energy from the renewable resources. The German Government's budget for energy research is reported €3.5 billion for the period 2011-2014[5]. Also, Germany gets 33 percent of its electricity from renewables energy sources. Alone in 2015, 193 billion kilowatt electricity has been produced from renewable energy sources. This makes Germany ranks second in the world in installed photovoltaic solar capacity.

Among the European countries, France shows tremendous research potential in areas related to electrochemical. In France, we identify clusters of scientific research related to Energy Storage and Green Chemistry.

In Canada, most prominent research topics in the field of renewable and sustainable energy are hydropower and bioenergy. According to International Energy Agency, Canadian renewable energy production is 18.9 percent of Canada's total primary energy supply[6]. Canada is also the second largest producer of hydroelectricity in the world.

5. Discussion and implications

This case study has several significant implications for human behavior analysis for library and information science. First, the use of advanced scientometrics and Data Mining methods in the context of scientific production and consumption reveals several patterns of human behavior. It is a critical input and a key requirement for understanding various aspects of know-how transfer at the international level. The adoption of advanced mining techniques to scientific data enables further insights as well for the design of sophisticated library systems. The integration of our approach with recommender systems technologies can justify sophisticated library services for the provision of the most relevant literature on demand. The use of text mining in relevance to the most influential aspect of scientific production can also support the design of advanced e-learning systems or active learning sessions for the promotion of research skills of students. The integration of advanced computational methods for Data Mining with cloud services can also deliver a marketplace of cloud-based open architectures for use in libraries. Distributed information systems and infrastructures for scientific knowledge management can also promote the vision of open access beyond specific publishers' systems. We do believe that this is a new exciting area for research. The detailed mapping of requirements is required for the development of flexible personalized services for the creation and consumption of scientific knowledge. Another, last but not least, area for future research is the human behavior analysis with focus on groups of researchers and influential personalities in scientific domains. The visioning of a Global Library Service beyond time and space limitations is a matter of time. From this point of view, additional research on the establishment of smart library interfaces is required.

6. Conclusions

In this paper, we have analyzed the production and consumption of knowledge flow originated from the scientific literature related to sustainable and renewable energy, during 1996-2009. We analyze the research topics produced by the USA that are consumed by the scientific community in different international contexts. For instance, Chinese scientific community heavily produces topics related to photovoltaic energy and biomass. While Japanese scientific community produces topics related to fuel cell, Taiwanese community shows remarkable competency in solar cells. Overall, we believe that such analyses may be helpful in establishing more effective multi-national research collaborations by understating the actual consumption of produced knowledge, not provided by studies of researcher mobility and co-authorship patterns. The contribution of our study in the special issue is related to the visioning of the application of advanced Data Mining methods for the delivery of sophisticated next-generation library services. The need to understand patterns of human behavior with the adoption of several emerging technologies is not just a necessity but a bold action of the global scientific community for the management of the collective wisdom, the promotion of enhanced collaboration and the creation of advanced systems to support individual groups and the global scientific worlds for prosperity, development and sustainability.

Notes

1. www.fs-unep-centre.org
2. www.worldenergy.org/
3. http://bbc.com/news/business-34527431
4. http://topics.amcham.com.tw/2015/09/is-renewable-energy-the-way-forward-for-taiwan/
5. http://ngm.nationalgeographic.com/2015/11/climate-change/germany-renewable-energy-revolution-text
6. www.nrcan.gc.ca/energy/renewable-electricity/7295

References

Alcacer, J. and Gittelman, M. (2006), "Patent citations as a measure of knowledge flows: the influence of examiner citations", *The Review of Economics and Statistics*, Vol. 88 No. 4, pp. 774-779.

Blei, D.M., Ng, A.Y. and Jordan, M. (2003), "Latent Dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3 Nos 4-5, pp. 993-1022.

Borgman, C.L. (1990), *Scholarly Communication and Bibliometrics*, University of California, Los Angeles, CA.

Cohen, W.M., Nelson, R.R. and Walsh, J.P. (2002), "Links and impacts: the influence of public research on industrial R&D", *Management Science*, Vol. 48 No. 1, pp. 1-23.

Ding, C.G., Hung, W.C., Lee, M.C. and Wang, H.J. (2017), "Exploring paper characteristics that facilitate the knowledge flow from science to technology", *Journal of Informetrics*, Vol. 11 No. 1, pp. 244-256.

Guevara, M.R., Hartmann, D., Aristarán, M., Mendoza, M. and Hidalgo, C.A. (2016), "The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations", *Scientometrics*, Vol. 109 No. 3, pp. 1695-1709.

Hassan, S.-U. and Haddawy, P. (2013), "Measuring international knowledge flows and scholarly impact of scientific research", *Scientometrics*, Vol. 94 No. 1, pp. 163-179.

Hassan, S.-U. and Haddawy, P. (2015a), "Tapping into scientific knowledge flows via semantic links", *15th International Conference in Scientometrics and Infometrics*, *Istanbul*.

Hassan, S.-U. and Haddawy, P. (2015b), "Analyzing knowledge flows of scientific literature through semantic links: a case study in the field of energy", *Scientometrics*, Vol. 103 No. 1, pp. 33-46.

Ingwersen, P.E.R., Larsen, B. and Wormell, I. (2000), "Applying diachronic citation analysis to ongoing research program evaluations", *Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, Information Today.

Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *The Quarterly Journal of Economics*, Vol. 108 No. 3, pp. 577-598.

Joint, N. (2008), "Bemused by bibliometrics: using citation analysis to evaluate research quality", *Library Review*, Vol. 57 No. 5, pp. 346-357.

Jones, B.F., Wuchty, S. and Uzzi, B. (2008), "Multi-university research teams: Shifting impact, geography, and stratification in science", *Science*, Vol. 322 No. 5905, pp. 1259-1262.

Khasseh, A.A., Soheili, F., Moghaddam, H.S. and Chelak, A.M. (2017), "Intellectual structure of knowledge in iMetrics: a co-word analysis", *Information Processing & Management*, Vol. 53 No. 3, pp. 705-720.

Lochan Jena, K., Swain, D.K. and Bihari Sahu, S. (2012), "Scholarly communication of the electronic library from 2003-2009: a bibliometric study", *The Electronic Library*, Vol. 30 No. 1, pp. 103-119.

Lockett, A. and McWilliams, A. (2005), "The balance of trade between disciplines do we effectively manage knowledge?", *Journal of Management Inquiry*, Vol. 14 No. 2, pp. 139-150.

Nagaoka, S. and Yamauchi, I. (2015), "The use of science for inventions and its identification: patent level evidence matched with survey", Research Institute of Economy, Trade and Industry (RIETI).

Nomaler, Ö. and Verspagen, B. (2008), "Knowledge flows, patent citations and the impact of science on technology", *Economic Systems Research*, Vol. 20 No. 4, pp. 339-366.

Porter, A.L. and Chubin, D.E. (1985), "An indicator of cross-disciplinary research", *Scientometrics*, Vol. 8 Nos 3-4, pp. 161-176.

Rinia, E.J., Van Leeuwen, T.N., Bruins, E.E., Van Vuren, H.G. and Van Raan, A.F. (2002), "Measuring knowledge transfer between fields of science", *Scientometrics*, Vol. 54 No. 3, pp. 347-362.

Stigler, S.M. (1994), "Citation patterns in the journals of statistics and probability", *Statistical Science*, Vol. 9 No. 1, pp. 94-108.

Trajtenberg, M. (1990), "A penny for your quotes: patent citations and the value of innovations", *The Rand Journal of Economics*, Vol. 21 No. 1, pp. 172-187.

Tsay, M.Y. (2013), "Knowledge input for the domain of information science: a bibliometric and citation analysis study", *Aslib Proceedings*, Vol. 65 No. 2, pp. 203-220.

Tsay, M.Y. and Shu, Z.Y. (2011), "Journal bibliometric analysis: a case study on the journal of documentation", *Journal of Documentation*, Vol. 67 No. 5, pp. 806-822.

Yan, E., Ding, Y. and Kong, X. (2012), "Monitoring knowledge flow through scholarly networks", *Proceedings of the American Society for Information Science and Technology*, Vol. 49 No. 1, pp. 1-5.

Yan, E., Ding, Y., Cronin, B. and Leydesdorff, L. (2013), "A bird's-eye view of scientific trading: dependency relations among fields of science", *Journal of Informetrics*, Vol. 7 No. 2, pp. 249-264.

Yuan, S. and Hua, W. (2011), "Scholarly impact measurements of LIS open access journals: based on citations and links", *The Electronic Library*, Vol. 29 No. 5, pp. 682-697.

Zahari, A.R. and Esa, E. (2016), "Motivation to adopt renewable energy among generation Y", *Procedia Economics and Finance*, Vol. 35, pp. 444-453.

Zhang, Q.P. (2013), "Characterizing scientific production and consumption in Physics", Scientific Reports No. 3.

Zhou, P.S. (2010), "A comparative study on communication structures of Chinese journals in the social sciences", *Journal of the American Society for Information Science and Technology*, Vol. 61 No. 7, pp. 1360-1376.

Zhuge, H. (2006), "Discovery of knowledge flow in science", *Communications of the ACM*, Vol. 49 No. 5, pp. 101-107.

Corresponding author
Miltiadis D. Lytras can be contacted at: mlytras@acg.edu