Guest editorial

1. Linked open data of bibliometric networks: analytics research for personalized library services

Introduction

The integration of smart library services in the urban context is a bold move toward sustainable and smart cities. The capacity of emerging technologies, such as data mining, advanced analytics, scientometrics, cognitive computing and immersive interfaces, to set up flexible, personalized spaces for content, context and community integration enhances the traditional value-adding models of libraries. In our perception, we are entering a new era of collective human wisdom, facilitated by intelligent library services promoting the role of libraries as hubs of collaboration and community enhancement.

The objective of the special issue is to communicate and disseminate recent computer engineering and library and information management research and application developments that demonstrate the capacity of emerging technologies to radically change the library experience . The thematic areas for the special issue focus are linked to the following areas of computational intelligence:

(1) Smart cities research for sustainable development.

(2) Community projects for smart library services in urban areas and cities.

(3) Personalized library services and advanced analytics including:

- social network analysis in publications;
- automatic semantic annotation of bibliographical resources;
- citation networks;
- rising stars analysis; and
- research matching.

(4) Integration of IoT, cognitive computing and cloud services.

(5) Design of context-aware, ubiquitous library interfaces.

(6) Promotion of international collaboration through knowledge management and know-how transfer.

The next generation library systems will face tremendous changes due to their integration with complementary domains and applications. From solutions related to unique learning experiences to fully functional mobile marketplaces of micro-content exchange and push content applications, a new era of library experience that will offer the potential to promote library services as transparent value-adding integrators has begun. In this context of smart cities-enabled library services, there are requirements for content development, new standards and innovative strategies and models for personalization, unique value proposition and innovations in use cases for new library services in large context.

Current applications of smart library services worldwide present a very interesting picture. Most of the dominant providers provide advanced content management services and context-aware solutions with limited integration to other domains. In parallel, innovative business models, methodologies and frameworks for smart cities promote the critical role of library services as boosters for the use and integration of scientific knowledge to innovation and training.

Personalized library services as an effective content management channel, initially converging with but then perhaps moving beyond social networks and mobile and ubiquitous technologies, offer added value to the library experience.

Several large-scale systems already provide a range of services to different stakeholders in the publishing and library domain, including publishers, editors, students, academia, training institutions and innovation centers. A key challenge for smart library systems to become more mainstream in this area is for traditional urban context to be critically enriched by technology-enabled components offering smart library experiences in the context of a greater vision for an International Global, Sustainable Smart Planet. The potential impact of analytics research for "Personalized library services: a smart cities primer" is reflected in surveys that indicate that the integration of data mining, IoT, cloud and cognitive computing research can radically change library business models, suggesting the capacity of the *Library Hi Tech* to redefine how knowledge, people and social infrastructures are integrated into the context of smart cities.

The purpose of the special issue is to present state-of-the-art approaches to, and examples of, advanced library systems and components for smart cities and urban applications. Manuscripts have been sought that address these areas. Novel approaches and sound technological solutions will be expected.

2. Unfolding the challenges for smart libraries

This edition presents key future dimensions that tap emerging tools and technologies including advance data mining and information retrieval techniques, scientometrics, social media analytics, cognitive computing for big data and deep learning, to set up flexible, personalized spaces for content, context and community integration to enhance the traditional value-adding models of libraries. In our perception, we are entering a new era of collective human wisdom, facilitated by intelligent library services promoting the role of libraries as hubs of collaboration and community enhancement.

The web is emerging as a preferred platform to publish open data and to interlink it together. This has result in the Web of Documents to be emerged as Web of Data. The emergence of the Web of Data is termed as linked open data (LOD), which consists of interlinked open data sets that are being published in different domains such as government, news, health and geography. Also, the trend of publishing LOD overcomes the limitations of publishing data from a bounded group of domain-specific data access boundaries to global data space. Even though the trend of publishing LOD has got potential in several domains such as government, music and health, but data from various domains such as digital libraries (and digital documents in general), scientific publications, archives and scholarly communications still need to be processed to make it part of global data space. If we consider only scientific publications/documents, the data about these millions of scientific documents that are published by different publishers are either untouched or a minor fraction of it is added in LOD cloud. The main reason is that data about scientific documents are published as a bounded group of publisher-specific templates. The same is true for data contained in digital documents, archives and scholarly contributions. This opens doors for new methodologies, processes, tools and framework to be developed which can deal with challenges of data extraction, processing, interlinking and publication of data from human to machine understandable format.

Given the better accessibilities of LOD and our abilities, as an information scientist, to process heterogeneous large-scale data sets more effectively, the field of library and information sciences (LIS) is currently at ripe for exploration. We believe that the future of LIS would stand on the following research agendas (see Figure 1): better understanding of social usage indices under the umbrella of Altmetircs to disseminate research outcomes; bibliometric network to mine complex networks including social media and LOD; information retrieval model by extracting textual and non-textual meta-data from full-text
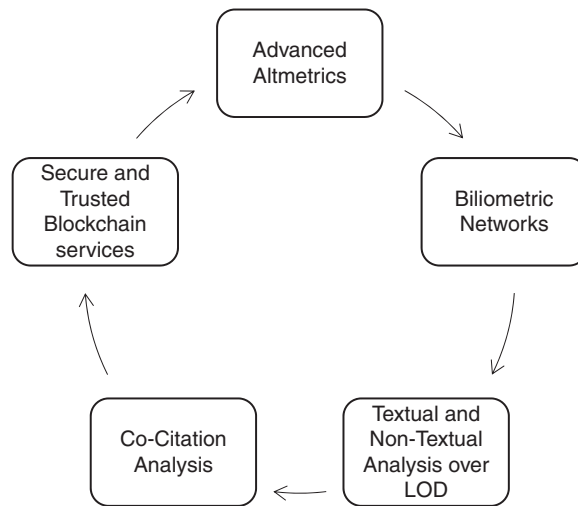
digital archives; improved methods of co-citation analysis, collaboration patterns, semantic analysis, etc. to support traditional bibliometrics; and last but not the least, exploiting the role of block chain-related technologies to ensure better data security and integrity of data used for evidence-based policy making.

First, under the umbrella of Altmetrics (Priem *et al.*, 2010), the scientific community would continue to explore the potential of social media contents and to understand better the power of evolving social media platforms. The studies have shown that arts and humanities researchers publish more diverse, non-article outputs than their peers in the sciences: books, book chapters, exhibition catalogs, textual data sets, book reviews, images, videos, audio files and many more. When it comes to research evaluation, humanities researchers are often not concerned with citations counts for their work. Instead, measures of impact in the humanities might include: prestige of publisher for one's book, peer-review status of work, prizes and awards received, mass media contributions or the number of national or international presentations completed.

The second key future direction would be the mining of complex networks generated by bibliography data set, social media connections or LOD (Jett *et al.*, 2017). Recent statistics reveal that bibliometric big data consist of over 20,000 journals along with over 0.9bn citations and 35m publications in Scopus. With the efforts of scientific community, a number of open source tools are now available to mine large-scale networks at different level of details such as citation network of publications/authors/journals; co-authorship network of authors/organizations; co-citation network of publications/authors/journals; bibliographic coupling network of publications/authors/journals or co-occurrence network of keywords/terms. For this analysis, the following open source tools are among the best choices and continuously improving: VOSviewer and CitNetExplorer (van Eck and Waltman, 2017), Gephi (Heymann, 2017) and Map Equation (www.mapequation.org). These tools have advanced visualization capabilities such as smart labeling algorithms, overlay visualization, built-in support for popular bibliography databases, text mining functionality, layout and clustering techniques. Overall, the advancements in mining large-scale networks would continue to increase for the better understanding of complex networks for informed decision making.

The third key direction of LIS would seek to provide better searching capabilities to users by coupling textual and non-textual meta-data from full-text digital archives (Cabanac *et al.*, 2018). Retrieval evaluations have shown that simple text-based retrieval methods scale up well but do

not progress. Traditional retrieval has reached a high level in terms of measures like precision and recall, but scientists and scholars still face challenges present since the early days of digital libraries. Thus, there is a sheer need in this direction, i.e., to identify relevant features from textual and non-textual meta-data items such as figures, tables, algorithms, etc., using unsupervised and supersized machine learning models, to build improved information retrieval techniques for digital libraries where simple text-based retrieval methods fail. Furthermore, another yet to be explored area of research is mining large-scale data sets such as digital libraries and books to create semantic abstractions for the better understanding of human-created knowledge.

Despite the fact that LIS community is leaping forward to explore new kind of data sources coupled with new tools and technologies, the traditional bibliometric analysis would keep playing a vital role to quantitatively assess the academic prestige of publication venues or author. We also expect to see more advanced bibliometric indices that utilize the advancements in context-based citation analysis, replacing traditional absolute citations-based indices such as impact factor or h-index. Moreover, we also foresee improved methods of co-citation analysis that taps the power of full-text for improved clustering accuracy of scientific publications.

Last but not the least, in the context of addressing basic contemporary societal concerns, such as transparency, accountability and trust in the policy-making process, LIS community would seek to exploit the role of blockchain technology in the near future to ensure better data security and integrity of data used for evidence-based policy making (Sicilia and Visvizi, 2019).

3. Overview of the published research in the special issue

This special issue publishes high-quality, original research contributions that address challenges in extraction, processing, interlinking and publishing data from different sources, specifically focusing on the data in the digital libraries, scientific publications, citation indexes, archives and scholarly communications that are jointly called bibliometrics networks. Undoubtedly, the role of a social media platform in relation to boost academic performance is vital with the emergence of Web 2.0 technology. However, social media platforms would vary across countries. Cheng *et al.* (2019) and Zheng *et al.* (2019) argue that in contrast to global acceptance of social media platforms, Chinese social media platforms are actually insufficient to support academic research, thus, make it difficult for the scholars to enhance the outreach of their scientific achievements.

Over decades, bibliometric studies that seek to investigate collaborative patterns have always been a point of great interest due to an implicit knowledge exchange between the collaboration. Cheng *et al.* (2019) and Zheng *et al.* (2019) show that more than half of the papers in the field of LIS, more specifically in *Library Hi Tech* journal between 2006 to 2017, have been written by more a single author. Similarly, Sabah *et al.* (2019) measure the effectiveness of scientific collaboration in terms of knowledge transfer from technologically advanced countries to developing counties. They present comprehensive use cases of top 50 research active Pakistani institutions in relation to their research outcomes with that of their international collaboration linkages with European and North American institutions. They show that a careful selection of collaborative partners can not only increase research output but it can significantly increase the quality of research.

Keeping in view that around 420m people in the world speak Arabic language, efforts need to be done in the context of building automated modeling techniques to improve digital access of scientific literature in under resource language like Arabic. An interesting work by Haraty and Nasrallah (2019), in this lieu, presents an auto-indexing modeling approach that employs association rule-based data mining techniques. Such models are extremely helpful to enhance the searching capabilities of search engines to facilitate user search needs.

Chui and Shen (2019) present tolerance analysis in scale-free social networks with varying degree exponents. Using the five state-of-the-art network indices, the authors

analyze conditions where the degree exponents may lie outside the range, i.e. [2, 3], in a scale-free network, named, average density, average clustering coefficient, average path length, average diameter and average node degree.

Finally, Sicilia and Visvizi (2019) propose the first blueprint of a form of sharing that complements open data practices with the decentralized approach of blockchain and decentralized file systems. They argue that the native decentralized (peer to peer) blockchain model offers encryption and validation that ensures the data not to be altered. Since the data are decentralized, cross-checked by the whole network and encrypted, it makes virtually impossible to hack data shared among the organizations to ensure better data security and integrity of data used for evidence-based policy making.

4. Conclusions

The evolution of smart services in modern libraries will continue to evolve. The technological enablers will be linked to the emerging technologies of artificial intelligence, cognitive computing, internet of things and analytics. Within this context a lot of work must be done in the policy making strategic level including:

- Design and implementation of smart cities and smart villages services to promote open access to knowledge and content (Visvizi and Lytras, 2018a, b).

- Integration of added to information systems in terms of perceived usefulness from real users (Lytras and Visvizi, 2018).

- Promotion of sophisticated technologies including cognitive computing and artificial intelligence, as carriers of advanced security, trust and ease of use (Lytras *et al.*, 2017).

- Understanding of advanced publication analytics in social impact terms. The quest of high-index values in current approaches, e.g., impact factor, h-index, must be rational and beyond manipulation strategies to promote the real social value of publications for sustainable socioeconomic growth.

We are happy to deliver this edition to the *Library High Tech* academic communities. We want to thank the editor in chief of LHT, Professor Michelle M. Kazmer, Florida State University School of Information: Florida's iSchool, USA, for the opportunity she offered us to serve the LHT community as well as the Emerald people for their continuous support and commitment in this project. Finally we are obliged to our distinguished contributors and reviewers for their excellent and high intellectual work. We do believe that the readers of LHT will value this special issue.

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

References

Cabanac, G., Frommholz, I. and Mayr, P. (2018), "Bibliometric-enhanced information retrieval: preface", *Scientometrics*, pp. 1-3.

Cheng, F.-F., Huang, Y.-W., Tsaih, D.-C. and Wu, C.-S. (2019), "Trend analysis of co-authorship network in Library Hi Tech", *Library Hi Tech*, Vol. 37 No. 1, pp. 43-56.

Chui, K.T. and Shen, C.-w. (2019), "Tolerance analysis in scale-free social networks with varying degree exponents", *Library Hi Tech*, Vol. 37 No. 1, pp. 57-71.

Haraty, R.A. and Nasrallah, R. (2019), "Indexing Arabic texts using association rule data mining", *Library Hi Tech*, Vol. 37 No. 1, pp. 101-117.

Heymann, S. (2017), "Gephi", *Encyclopedia of Social Network Analysis and Mining*, pp. 1-14.

Jett, J., Cole, T.W., Han, M.J.K. and Szylowicz, C. (2017), "Linked Open Data (LOD) for library special collections", *Proceedings of the 17th ACM/IEEE Joint Conference on Digital Libraries, IEEE Press*, pp. 309-310.

Lytras, M.D. and Visvizi, A. (2018), "Who uses smart city services and what to make of it: toward interdisciplinary smart cities research", *Sustainability*, Vol. 10.

Lytras, M.D., Raghavan, V. and Damiani, E. (2017), "Big Data and data analytics research: from metaphors to value space for collective wisdom in human decision making and smart machines", *International Journal on Semantic Web and Information Systems*, Vol. 13 No. 1, pp. 1-10.

Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2010), "Altmetrics: a manifesto".

Sabah, F., Hassan, S.-U., Muazzam, A., Iqbal, S., Soroya, S.H. and Sarwar, R. (2019), "Scientific collaborations networks in Pakistan and their impact on institutional research performance: a case study based on scopus publications", *Library Hi Tech*, Vol. 37 No. 1, pp. 19-29.

Sicilia, M.-A. and Visvizi, A. (2019), "Blockchain and OECD data repositories: opportunities and policymaking implications", *Library Hi Tech*, Vol. 37 No. 1, pp. 30-42.

van Eck, N.J. and Waltman, L. (2017), "Citation-based clustering of publications using CitNetExplorer and VOSviewer", *Scientometrics*, Vol. 111 No. 2, pp. 1053-1070.

Visvizi, A. and Lytras, M.D. (2018a), "Editorial: policy making for smart cities: innovation and social inclusive economic growth for sustainability", *The Journal of Science and Technology Policy Management*, Vol. 9, pp. 1-10.

Visvizi, A. and Lytras, M.D. (2018b), "Rescaling and refocusing smart cities research: From mega cities to smart villages", *The Journal of Science and Technology Policy Management*, Vol. 9, pp. 134-145.

Zheng, W., Wu, Y.J. and Lv, Y. (2019), "More descriptive norms, fewer diversions: boosting Chinese researcher performance through social media", *Library Hi Tech*, Vol. 37 No. 1, pp. 72-87.

Further reading

Daud, A., Amjad, T., Siddiqui, M.A., Aljohani, N.R., Abbasi, R.A. and Aslam, M.A. (2019), "Correlational analysis of topic specificity and citations count of publication venues", *Library Hi Tech*, Vol. 37 No. 1, pp. 8-18.

Visvizi, A., Mazzucelli, C. and Lytras, M., "Irregular migratory flows: towards an ICT" enabled integrated framework for resilient urban systems", *The Journal of Science and Technology Policy Management*, Vol. 8, pp. 227-242.

7