Journal of Intelligence Studies in Business

Journal of Intelligence Studies in Business

Publication details, including instructions for authors and subscription information: https://ojs.hh.se/index.php/JISIB/index

The potential of business intelligence tools for expert finding

Mehdi Dadkhah[a], Mohammad Lagzian[a]*, Fariborz Rahimnia[a], Khalil Kimiafar[b]

[a]Department of Management, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Iran; [b]Department of Medical Records and Health Information Technology, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran

*m-lagzian@um.ac.ir

To cite this article: Dadkhah, M., Lagzian, M., Rahimnia, F. & Kimiafar, K. (2019) The potential of business intelligence tools for expert finding. *Journal of Intelligence Studies in Business*. 9(2) 82-95.

Article URL: https://ojs.hh.se/index.php/JISIB/article/view/411

Journal of Intelligence Studies in Business
Vol. 9, No. 2 (2019) pp. 82-95
Open Access: Freely available at: https://ojs.hh.se/

Stevan Dedijer
1911-2004

The potential of business intelligence tools for expert finding

Mehdi Dadkhah[a], Mohammad Lagzian[a]*, Fariborz Rahimnia[a] and Khalil Kimiafar[b]

[a]*Department of Management, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Iran*
[b]*Department of Medical Records and Health Information Technology, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran*

Corresponding author (*): m-lagzian@um.ac.ir

Received 30 July 2019 Accepted 28 October 2019

ABSTRACT Finding the right experts for data gathering through interview serves as a key for particular research works. However, most expert finding methods in the literature require great deals of technical knowledge, making them somewhat impracticable for business researchers without deep technical knowledge. Accordingly, there is a need for an expert finding solution for researchers without a deep technical background. As business researchers may have knowledge about business intelligence and its tools, the use of business intelligence tools can be used to solve such issue. The present paper discusses the process of using business intelligence tools to find potential experts for example topics. Subsequently, based on a literature review, criteria are presented for distinguishing different experts. Finally, the analytic hierarchy process is discussed for assigning weights to both selection criteria and potential experts. The audience of this paper is researchers who are familiar with business intelligence tools or would like to learn how to work with them.

KEYWORDS Business intelligence, business intelligence tools, expert selection, expert selection criteria, participant selection

1. INTRODUCTION

In social science, qualitative methods are popular for conducting research. In the qualitative research methods, interviews with participants are one data collection instrument (Louise Barriball & While, 1994). Accordingly, different strategies are presented for selecting potential participants. In some cases, unavailability of participants for face-to-face interviews or other difficulties led researchers to utilize computers as a research instrument (Girvan & Savage, 2013; Markham, 2004). A review of the literature on research methodologies shows that, unlike quantitative research, qualitative research tends to select participants purposively (Flick, 2008; Marshall, 1996) based on specific criteria. In such studies, the researcher decides, based on the specific criteria, who to consider as a participant for the research (Flick, 2008; Marshall, 1996). There are a number of strategies for purposive sampling in qualitative research (Palinkas *et al.*, 2015). As described by Palinkas *et al.*, such strategies can be grouped into three major categories: (1) the strategies emphasizing similarity, (2) the strategies emphasizing variation, (3) and the strategies with no specific emphasis (Palinkas *et al.*, 2015; Patton, 2002). Despite the apparently extensive research on purposive sampling in qualitative research, it is not always an easy task to accomplish. It is not always easy to find participants for a research plan where the required data shall be obtained from people with professional knowledge (i.e. experts). The situation becomes even more critical when such expert experience falls within multiple contexts, with only few experts in each context, or when such experts are in multiple locations (for example, country, university, or organization) making it impossible for the researcher to become aware of all of them. Even though snowball sampling can be a good alternative for such conditions, finding participants within a reasonably short period of time is also an issue. Finding expert participants for a qualitative research may be difficult in some cases. This problem is not limited to some cases in qualitative research: there are studies that discuss this issue without considering this domain (Gretsch, Mandl, & Hense, 2011; Ru, Xu, & Guo, 2007; Serdyukov & Hiemstra, 2008). So, finding the right expert can be a challenging task. In such situation, using a machine-made method for finding experts can be helpful.

The present research aims to show how to use business intelligence (BI) tools and the analytic hierarchy process (AHP) to find experts. The target audience of the current study is researchers who are interested in BI or have knowledge in this regard. For example, business students can learn to work with the tools used in this paper as they may learn BI in university or at workshops, Section 2 gives a brief overview of some available methods. Section 3 describes the process of using BI tools to find experts and presents a discussion on its results. The final conclusions are drawn in Section 4.

2. BRIEF OVERVIEW OF EXPERT FINDING RESEARCH

Researchers have presented various methods to find experts. Deng *et al.* presented three models for finding experts by using DBLP bibliography and Google Scholar services (Deng, King, & Lyu, 2008). Naeem *et al.* utilized data mining for the same purpose (Naeem, Khan, & Afzal, 2013). Kardan *et al.* presented and discussed a model for expert selection in social networks (Kardan, Omidvar, & Farahmandnia, 2011). Other research focuses on finding experts in social networks or community question answering websites (Bozzon, Brambilla, Ceri, Silvestri, & Vesci, 2013; Kao, Liu, & Wang, 2010; Kardan et al., 2011; Riahi, Zolaktaf, Shafiei, & Milios, 2012; Zhang, Tang, & Li, 2007; Zhao, Zhang, He, & Ng, 2014). Wang *et al.* proposed an algorithm, called *ExpertRank*, that identifys and evaluates experts based on both documentation and an individual's authority in his or her knowledge community. This algorithm is a modification of the *PageRank* algorithm to evaluate an individual's authority (Wang, Jiao, Abrahams, Fan, & Zhang, 2013). Demartini used Wikipedia as the knowledge source to find experts in topics. He used *WordNet* and *Yago* to improve retrieval effectiveness (Demartini, 2007). Zhan *et al.* employed probabilistic latent semantic analysis to propose a mixture model for expert finding. Semantic themes will be identified by such mixture models between terms and documents. Then by using these themes, their method finds relevant experts based on the query (Zhang, Tang, Liu, & Li, 2008). Yang *et al.* proposed an expert finding system by analyzing an individual's journal papers. They state that journal publication can be used to find the expertise of a researcher (Yang, Chen, Lee, & Ho, 2008). Lin *et al.* in a survey discussed methods and models that focus on expert findings and show the current status of research in this regard (Lin, Hong, Wang, & Li, 2017). Boeva *et al.* proposed a data driven expert finding technique. Their technique also weighs experts based on their expertise (Boeva, Angelova, & Tsiporkova, 2017). Further search into the literature would highlight other technical methods for expert finding.

Even though these are valuable and interesting, such methods are only useful for researchers with advanced technical knowledge. Other researchers without deep technical knowledge may not be able to take advantage of such techniques, unless the technical methods are translated into convenient tools for social science researchers. There are some easy-to-use expert finding methods in the literature. On its user interface, Scopus provides an interested option for analyzing search results (Beatty, 2015), offering an easy-to-use method for non-technical researchers who are looking for particular experts. This method can be used for expert selection. Schuemie and Kors developed a web-based tool entitled *Jane* (http://jane.biosemantics.org/index.php) which can be used for expert finding. *Jane* uses PubMed as the source of data and presents result by using the Lucene MoreLikeThis algorithm and k-nearest neighbor approach (Schuemie & Kors, 2008). Cifariello *et al.* developed a semantic search engine entitled

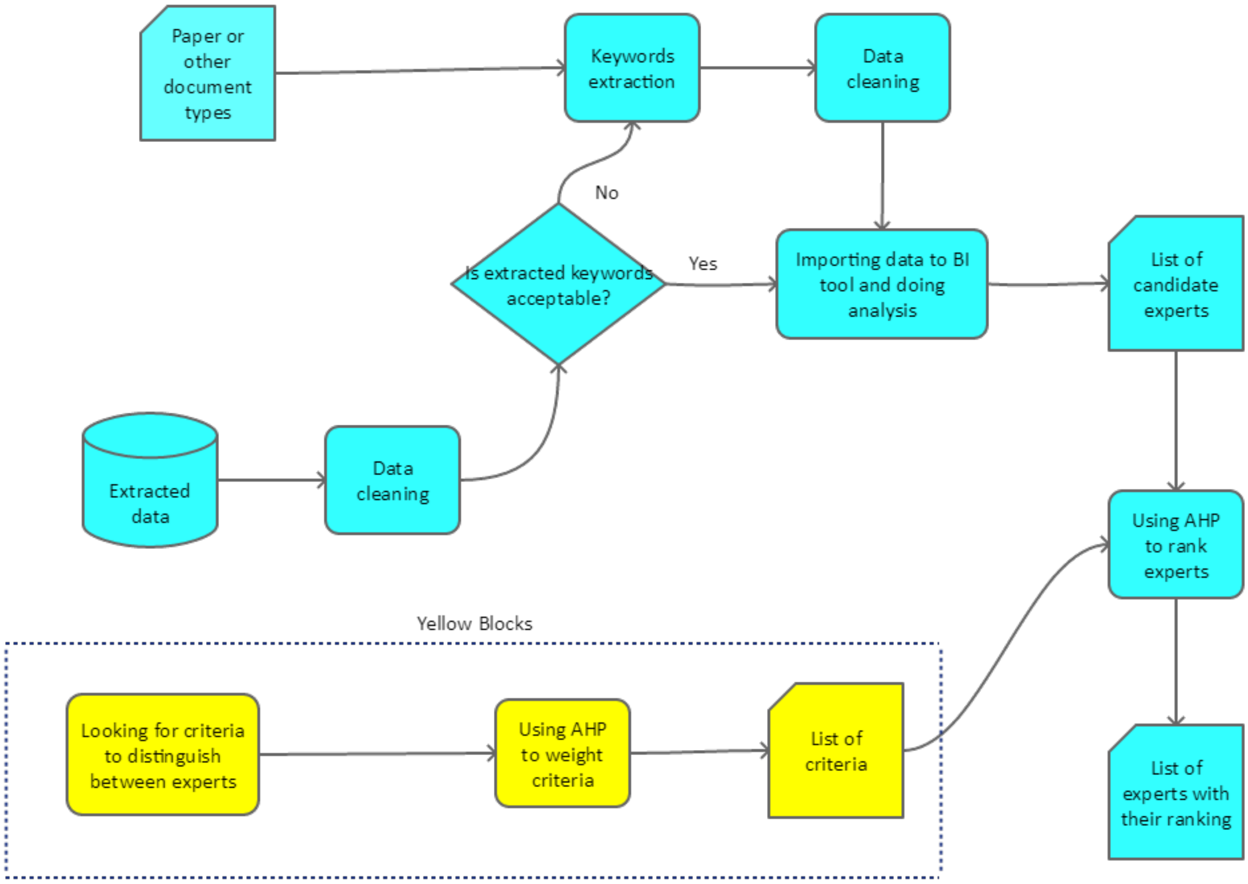

Figure 1 Schematic presentation of the proposed method for expert selection.

Wiser that finds experts. It models each author's expertise with a graph by using Wikipedia. Experts are identified through co-occurrence of searched keywords in their publications and this graph. Wiser has an online graphical-based version (https://wiser1.sobigdata.d4science.org/search) which works based on University of Pisa publications (Cifariello, Ferragina, & Ponza, 2019). These tools help researchers to find experts. However, when a user with no advanced technical knowledge aims to analyses his or her own data or data related to other academic sources, this method is not helpful. It should be noted that is possible to the adapt proposed method in literature to be used for different data sources, but technical knowledge in this regard is required. BI tools are especially useful for business students to find experts. This study focusses on a process that helps researchers to find experts by utilizing BI tools. The process in this paper can be used by individual who are familiar with BI to find experts. This paper does not present a new method, it shows the capability of existing BI tools to be used for expert finding.

3. PROCESS OF FINDING EXPERTS USING BI TOOLS

Today, organizations are encountering large sets of data that cannot be used without BI. In order to make better decisions, organizations utilize BI to create knowledge out of their data (Chaudhuri, Dayal, & Narasayya, 2011). A BI solution follows a BI architecture. Generally, companies store different data of different sources. However, before a BI solution can be successfully implemented, the entire set of such data must be integrated to a data warehouse by using a special process called ETL (extract, transform and load). Given the inefficiency of executing queries on an entire set of data in an organization, it is necessary to extract related data before proceeding to executing such a query. Once an integrated data warehouse is developed, different servers can efficiently access the data in the warehouse through front-end applications. Such an application can be used by particular decision-makers depending on their roles in the organization (Chaudhuri *et al.*, 2011; Negash, 2004; Sherman, 2014). Details of BI are out of scope of the present work, where only the BI tool is used, rather than a full BI implementation. A BI tool is a vendor's software that is used to develop BI applications or styles (e.g. dashboards or scorecards) (Sherman, 2014).

Table 1 The data extracted from Scopus by searching the term "Internet of Things".

Author Ids	Title	Year	Source title	Author Keywords
Records of data				
14018777000; 27867946500; 57202208939; 57202211443; 38461465700;	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
57202334348;	FAN: Framework for authentication of nodes in mobile ad hoc environment of internet-of-things	2019	Advances in Intelligent Systems and Computing	Access control; Internet-of-Things; Mobile ad hoc network; Secure permission; Security; Ubiquitous
57203555315; 56238720400;	Study and design of smart embedded system for smart city using internet of things	2019	Lecture Notes in Electrical Engineering	Electronic devices; Internet of Things (IoT); Smart city
Other records of data				

Table 2 The cleaned data for the analysis in this study.

Author Ids	Title	Year	Source title	Author Keywords
Records of data				
14018777000	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
27867946500	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
57202208939	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
57202211443	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
38461465700	Multidimensional wavelet neuron for pattern recognition tasks in the internet of things applications	2019	Advances in Intelligent Systems and Computing	Classification; Internet of things; Machine learning; Multidimensional wavelet neuron; Online learning; Pattern recognition
57202334348	FAN: Framework for authentication of nodes in mobile ad hoc environment of internet-of-things	2019	Advances in Intelligent Systems and Computing	Access control; Internet-of-Things; Mobile ad hoc network; Secure permission; Security; Ubiquitous
57203555315	Study and design of smart embedded system for smart city using internet of things	2019	Lecture Notes in Electrical Engineering	Electronic devices; Internet of Things (IoT); Smart city
56238720400	Study and design of smart embedded system for smart city using internet of things	2019	Lecture Notes in Electrical Engineering	Electronic devices; Internet of Things (IoT); Smart city
Other records of data				

Partially inspired by the general BI solution, and its uses for academic research introduced by Chaudhuri et al. (2011), Sherman (2014) and Dadkhah and Lagzian (2018) the process of experts finding is schematically presented in Figure 1. Similar to the work by Boeva *et al.*, the process herein uses a keyword-based search to identify experts (Boeva, Angelova, & Tsiporkova, 2017). A basic requirement of a BI process is data. The data may come from different sources. In the field of research, such data may be collected from academic databases such as Scopus or Google Scholar, academic papers, un-published documents, or reports. For the most part, the academic databases provide the user with an option to extract relevant data based on various criteria. For example, upon searching Scopus for the term "Internet of Things", one can extract the *titles*, *authors' names*, *keywords*, and/or *names of the journals* corresponding to the search, resulting in a file of a particular format. Figure 1 highlights such data as "extracted data". When it comes to possibly large offline documents on a local disk, there is a need for methods to either automatically extract such data and print that into a file or do the same manually. Various methods have been proposed for keyword extraction in the literature (MATSUO & ISHIZUKA, 2004; Merrouni, Frikh, & Ouhbi, 2016; Rose, Engel, Cramer, & Cowley, 2010). In such processes, keywords play a fundamental role. The present work is focused on two features in each document: the author's name and keywords. Table 1 shows a summary of the data extracted from Scopus by searching the term "Internet of Things", as an example. This search was limited to 2000 records by the authors (search date: 7 September 2018). Accordingly, the following features were included in the data: Author Id, Title, Year, Source title, Author Keywords.

Upon extracting the relevant data, one should check for possible inconsistencies, errors or related issues. For example, there may be duplicate records to be cleaned up or inconsistencies to be addressed by reformatting the data. The data cleanup stage is critical for the successful accomplishment of the entire process. In the present work, an easy-to-use freeware called OpenRefine was used to clean up the data ("OpenRefine," 2018) (Verborgh & De Wilde, 2013). After the cleanup stage, one should evaluate the acceptability of the extracted keywords. If the keywords were found to be unacceptable, automatic keyword extraction methods can be applied to extract other keywords. Table 2 shows the extracted data following the cleanup stage. As suggested by the designation, *Author Id*s indicate the authors' names and help classify keywords by authors. Accordingly, a single *Author Id* was presented per row. Also, correction may be necessary for spelling multiplicity in the source title. The records lacking an *Author Id*, with the corresponding field left blank, were deleted in this study. In Table 2, each row refers to a particular author and provides details of paper title, year of publication, place of publication, and keywords.

At this stage, the dataset is ready for analysis. This paper deals only with the BI tool rather than a full BI implementation. There are different BI tools with different features, and their associated costs vary from free to paid. BI tools provide different features including dashboards and reporting capability. Dashboards provide graphical elements for data visualization. Reporting capability lets the user use the information element (Bernardino & Tereso, 2013). Both reporting and dashboard elements can be used to find relevant experts. In this paper, a trial licensed version of DBxtra (https://dbxtra.com) was used as we had access to it, and it provided a drag-and-drop option. The documentation of this tool provides a good source for operating the software (DBxtra, 2018). Utilizing the software, a constraint was set to consider only records for which at least two features were available: author's name and keywords. Then the authors were filtered based on keywords to find relevant experts. This is why the present method was said to be based on keywords. For example, we filtered authors by selecting the keywords "energy", "sensor" and "IoT", then the software listed the authors who published papers contained these terms as keywords. In DBxtra, a dashboard is designed using a list box, two combo boxes, a chart, and a pivot table. For the example considered in this research, the list box contained the *Author Keywords* values. Accordingly, a list of relevant experts could be obtained by applying a filter on this list. As shown in Figure 2, a filter was designed to extract the list of authors who had used the terms "energy", "sensor" and "IoT" as keyword.

The two combo boxes could filter the data by year and place of publication, with the chart indicating the count of candidate experts.

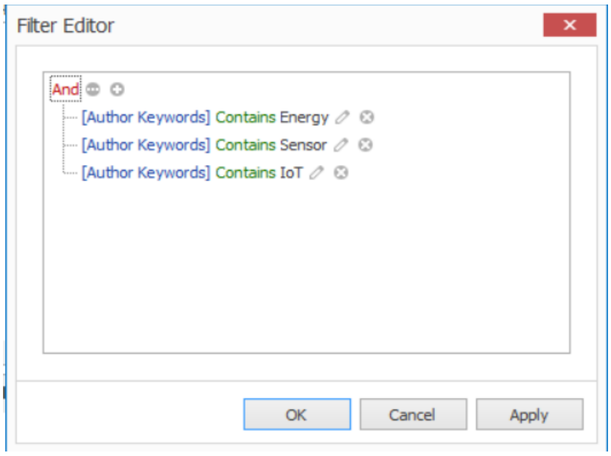

Figure 2 The filter applied on the list box to extract the list of authors who had used the terms "energy", "sensor" and "IoT" as keyword.

Figure 3 shows the dashboard with the experts who had paper(s) containing the following keywords: "energy", "sensor" and "IoT". Accordingly, a list of 66 experts with expertise related to sensors and energy in the IoT domain was obtained. There are different dashboard elements that researchers can refer to in order to document and understand their tools. By using such elements, there is the possibility to visualize data and do relevant analyses, then find experts. When the data is clean, the availability of working with BI tools and their elements plays an important role in finding relevant experts from data. Based on their needs, researchers should decide which elements are helpful for their analysis add them to their dashboard. Also, each element needs to be configured. For example, the chart in Figure 3 is configure to count the number of Author IDs in the data. Generally, it counts a distinct value of Author IDs in all data. The combo box is configured to include data related to the keywords. When a filter is applied on this combo box, the chart counts only the Author IDs that are accessible through this filter. We do not discuss more about the capabilities of each BI tool and their related elements, as there is good documentation in this regard.

3.1 Ranking experts based on the research topic

Once one is finished identifying the relevant experts, it is possible to evaluate the suitability of such potential experts for the research. The BI tool provides potential experts and next the researchers should confirm result. They should evaluate each potential expert to understand if the person is a suitable expert. As an example, one may need only 10 experts. If the BI tool provided 66 experts (Figure 3), one must select the 10 most suitable experts. For this purpose, beginning with an attempt to distinguish between experts based on some general criteria, one should remember that specific research exists with additional features for the purpose. In this study, relevant features were

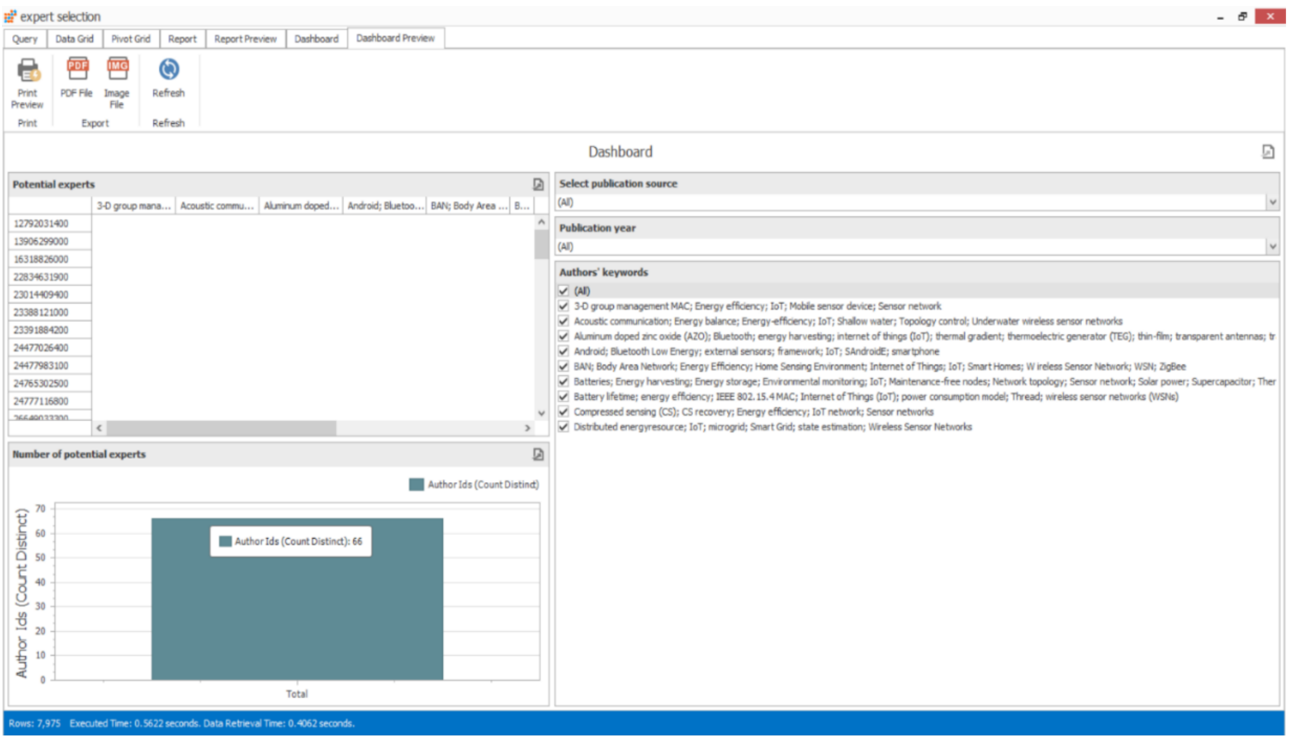

Figure 3 The dashboard designed for finding experts who had published paper(s) containing the following keywords: "energy", "sensor" and "IoT".

identified by looking into the literature. Accordingly, papers presenting criteria for expert selection were identified (Afzal, Kulathuramaiyer, & Maurer, 2008; Afzal & Maurer, 2011; Benner, Tanner, & Chesla, 1992; Boeva et al., 2017; Cameron, Aleman-Meza, Decker, & Arpinar, 2007; Hirsch, 2005; Naeem et al., 2013; Quatrini Carvalho Passos Guimarães, Pena, Lopes, Lopes, & Bottura Leite de Barros, 2016); (Academia Europaea as cited in Naeem *et al*., 2013; Pakistan Academy of Sciences as cited in Naeem *et al*., 2013; Fehring as cited in Quatrini Carvalho Passos Guimarães, Pena, Lopes, Lopes, & Bottura Leite de Barros, 2016). As some of these papers were subject-oriented, respective criteria were generalized and used as a feature for expert identification and ranking (Table 3). Researchers may need to define new criteria based on their research.

In the next step, an analytic hierarchy process (AHP) can be used to assign weights to the features to facilitate the process of decision-making for expert selection. AHP refers to a pairwise comparison method for weighting a pool of alternatives, so as to select an alternative based on particular criteria. Using multilevel hierarchic structures, an AHP involves alternatives, criteria, and a goal. It has been widely used in business- and government-led applications (Saaty, 1977, 1990, 2013). In this paper, AHP is utilized to rank a set of candidate experts based on particular criteria extracted from the literature, for the purpose of final expert selection. AHP arranges the decision criteria into a hierarchical structure. In this stage, the scale shown in Table 4 can be used as a foundation to design a questionnaire for pairwise comparison (Saaty, 1977, 1990, 2013).

Table 3 The features used for selecting and ranking the experts (adapted from the references cited in the text).

No.	Feature	Description
1	Projects	To distinguish experts participating in a particular project(s).
2	Awards	To distinguish experts who have achieved a particular award(s)
3	Honorarium	To distinguish experts who have contributed into a particular domain(s).
4	Affiliations	To distinguish experts with a particular affiliation(s), taking the affiliation as a measure of proficiency in a particular domain(s).
5	Request for Comments (RFC)	To distinguish experts who were frequently requested for comments, taking RFC as a measure of experimental skills in a particular domain(s).
6	Supervision	To distinguish experts who are active in the field of academic supervision of students.
7	Collaboration	To distinguish experts who have collaborated with others at international level.
8	Relevance	To distinguish experts who are actually relevant to the considered research.
9	Keynote Speaker	To distinguish experts who have been a keynote speaker in a conferences or other societies.
10	Reviewer	To distinguish experts with the required deals of skill and expertise to serve as a reviewer for a journal or conference.
11	Protocol Design	To distinguish experts with the required deals of skill and knowledge to design protocol standard(s).
12	Distinctions	To distinguish outstanding experts, in comparison to peers.
13	Citation number	To distinguish experts with a particular number of received citations.
14	Publication number	To distinguish experts with a particular number of publications.
15	Co-author network	To distinguish experts who have worked with a particular number of co-authors.
16	Academic degree	To distinguish experts with a particular academic degree.
17	Gender	To distinguish experts of a specific gender.
18	Experience duration	To distinguish experts with a particular number of years of contribution into the considered domain.
19	Extent of citations in given domain	To distinguish experts based on the number of received citations in a particular domain: $Extent\ of\ Citation = \frac{total\ number\ of\ received\ citation\ in\ a\ topic\ by\ candidate\ expert}{total\ number\ of\ received\ citation\ in\ a\ topic}$
20	Impact factor of publication journals	To distinguish experts who had papers published in journals of particular impact factor(s).
21	H-Index	To distinguish experts based on the metric proposed by Hirsch. This metric indicates the j number of papers that received j or higher number of citations.
22	Researcher profile	To distinguish experts based on their profile in terms of relevant skills, keywords, and topics of interest.

Table 4 Scales for comparing alternative experts (Saaty, 1977, 1990, 2013).

Numeric scale	Meaning
1	The two alternatives are equally important.
3	An alternative is moderately more important than another.
5	An alternative is essentially more important than another.
7	An alternative is strongly more important than another.
9	An alternative is extremely more important than another.
2, 4, 6, 8	Intermediate values between the above milestones.

The yellow blocks in Figure 1 show the corresponding steps through the whole process. If there is uncertainty in decision making, fuzzy AHP can be used. It uses fuzzy numbers as the numerical scales (Özdağoğlu and Özdağoğlu, 2007; Wang and Chin, 2011; Ramík and Korviny, 2010). the value of features in Table 3 should be gathered or calculate manually for each candidate expert, but it is possible to use a programming language to automate some tasks.

In order to implement AHP in this study, the experts list and the features described in Table 3 were taken as the alternatives and criteria, respectively (Figure 4). Then, two pairwise comparison questionnaires can be designed for the considered criteria and experts. The questionnaires should be presented to a number of university professors and researchers in the field of research. The data extracted from the questionnaires can be analyzed using different tools such *Super Decisions*, a tool for multi-criteria decision making (SuperDecsion, 2018), and the results should be used to assign weights to the criteria and experts. Researchers usually need to select and rank such criteria for their research activities, as may be necessary depending on the specific research question(s). Figure 4 shows the hierarchy of the AHP model developed for expert selection. In this study, weights are calculated for each criterion and the BI tool provides a list of potential experts as alternatives for this model (Figure 4). In a final step, the experts were ranked based on the criteria.

In AHP, the consistency ratio shall be equal to or smaller than 0.1; otherwise the result of pairwise comparison may be unreliable (Saaty, 1977, 1990, 2013). Indeed, the consistency ratio increases when increasing the number of elements in a comparison (Benítez *et al.*, 2011). Accordingly, the use of 22 criteria or an expert list with many candidates in the developed AHP model may lead to a consistency ratio exceeding 0.1, indicating unreliable results. However, researchers could choose to select only a subset of the 22 criteria, depending on the scope of their research, or narrow their queries to find a smaller number of experts. Other methods have also been proposed for addressing the problem of inconsistency in AHP (Benítez *et al.*, 2011; Benítez *et al.*, 2012). The value of features shown in Table 3 should be determined manually by researchers, however it is possible to gather values for some of the features automatically. For example, h-index and total citation count, number of

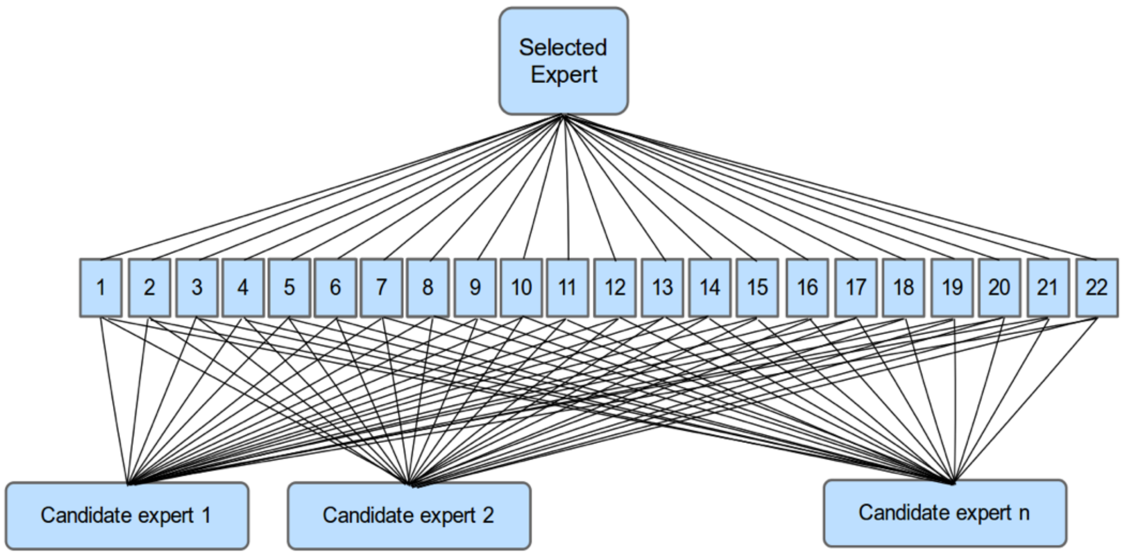

Figure 4 The AHP model developed for expert selection.

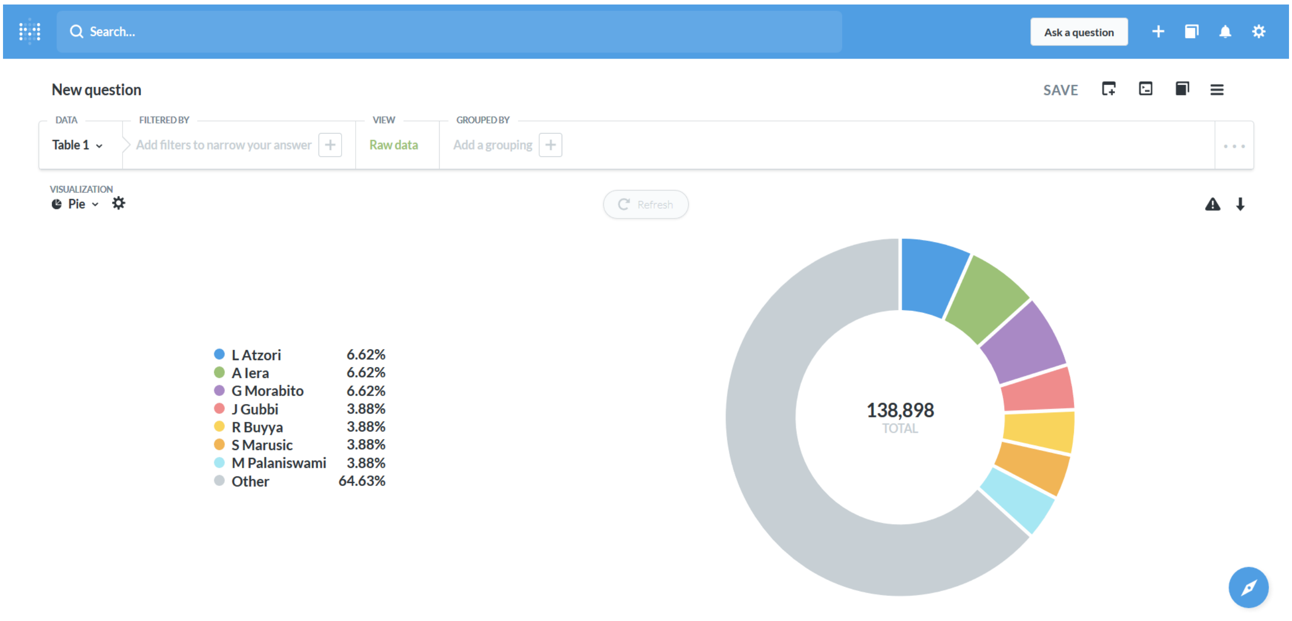

Figure 5 Identified experts using Metabase.

publication and co-authors can be gathered through Scopus.

By using BI tools, it is possible for researchers to do more advanced analysis on their data. For example, by extracting data from Scopus by searching the term "Internet of Things", it is possible to find experts with different conditions such as:

> *Experts who published a paper about the Internet of Things AND started their publication in this topic at least 5 years ago AND have a total citation count on this topic above 1200 AND have the article type "Journal Paper" AND are affiliated to a specific country AND published by a specific publisher AND published in a top information system journal.*

It is possible to add four columns including journal impact factor, author h-index, publication number, and total number of co-authors to the extracted data from Scopus. Here we attach new data to the extracted data from Scopus. This data is the value of the four features discussed in Table 3. Now, the previous query could be more advanced as:

> *Experts who published a paper about the Internet of Things AND started their publication in this topic at least 5 years ago AND have citation counts on this topic higher than 1200 AND their article type is a Journal Paper AND are affiliated to a specific country AND published by a specific publisher AND published in a top information system journal AND published in a journal with an IF higher than 1 AND with a total number of published papers higher than 10 AND author's h-index is higher than 5 AND total number of co-author is higher than 12*

This process can be done through other tools and data sources. To evaluate this expert finding process, researchers used two other tools and tried to find potential experts who are familiar with both the internet of things (IoT) and patient monitoring. The researchers are interested in experts who received at least 700 citations on a publication in this topic and published it at least five years ago. They used Publish or Perish (Publish or Perish, 2018) to extract data from Google Scholar and Metabase (Metabase, 2018) to analyses the data (search date: 10 November 2018). As Publish or Perish does not provide keywords for each paper, there are two option to find keywords: 1) use methods for extracting keywords from papers, 2) narrow the search by defining all keywords then analyzing the result instead of doing a broad search and then limiting result by keywords. Figure 5 shows the output of the analysis in Metabase. Based on this analysis, the researchers found 15 potential experts. In the extracted data from Google scholar via Publish or Perish, there are other features including Source Title, Publisher, Article URL, Cites Per Year, Author Count, and Title of Papers. This means that it is possible to use these features to do more advanced searches to find potential experts

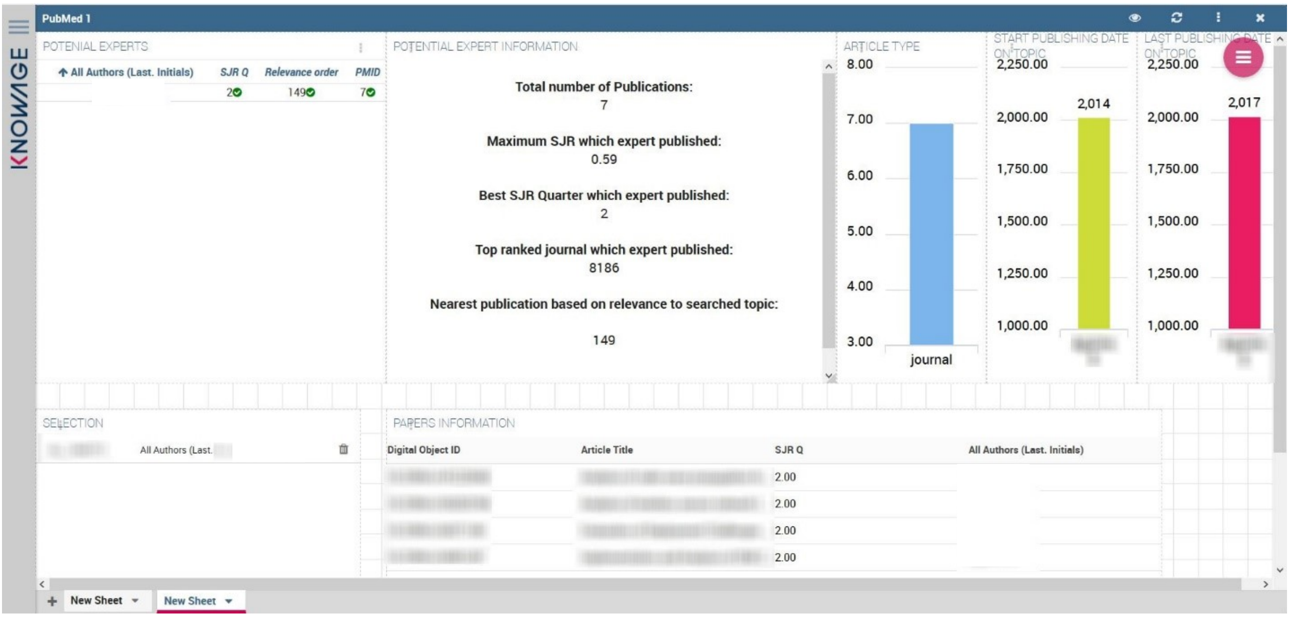

Figure 6 A dashboard in Knowage for finding experts. In this dashboard, an expert has been selected and the information is shown.

from this data. After finding expert via the BI tool, now we can manually review experts and use the Table 3 criteria to confirm experts with regard to our research.

Experts can also be found via PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) as the source of data, and Knowage (https://www.knowage-suite.com/site/home/) as the business intelligence tool. We searched for *"wireless sensor network"* (search date: 24 July 2019) and download all 769 result as the XML file. By using PubMed2XL (available from http://blog.humaneguitarist.org/projects/pubmed2xl/), the XML file was converted to an Excel spread sheet (Isaak, 2016). By using OpenRefine, the data was clean. As with *Jane*, it is possible to find potential experts based on the relevance of keywords. In PubMed, data is sorted according to its relevance to the search term, then downloaded. By having relevance of data to searched terms, it is possible to find experts based on relevance. By doing this, it is concluded that from the top 20 identified potential experts in Knowage, 15 of them were also in the list of retrieved experts from *Jane*. The difference was their rank compare to the *Jane* result. It is possible to get a list of potential experts who are familiar with wireless sensor networks by using the extracted data from PubMed. For example, we can find all individuals who have at least four publications about wireless sensor networks and at least one publication in the top 300 results, based on relevance. It is possible to do a more advance query to find individuals who have at least four publications about wireless sensor networks and at least one publication in the top 300 results based on relevance and at least one publication published in a journal in the first two quarters of the Scimago journal ranking (SJR). This query needs to merge new data with extracted data from PubMed. The *SJR* data can be retrieved from the SCiMago journal ranking website (available from https://www.scimagojr.com/journalrank.php). Then it is possible to merge the data together by using available tools such as OpenRefine. Figure 6 illustrates the dashboard in Knowage for finding such experts. This dashboard also has extra filters to find experts such as the first year of publication and the number of publications in a top quarter journal. It also shows some information about experts and their relevant papers.

The process discussed in this paper was also tested to find research method experts from a personal repository, and another study about knowledge management. Using this method was helpful for both this study and to simplify the expert finding task. In the earlier expert finding task, eight potential experts of the former 10 potential experts were identified. The main advantage of the process compared to most expert finding methods is that it has lower requirements for individual BI tool technical knowledge. BI tools currently support different options (for example drag and drop) to simplify the data analysis task (Smuts, Scholtz, & Calitz, 2015). By using a BI self-service tool, individuals can use BI tools with less technical knowledge (Imhoff & White, 2011).

Table 5 Comparison between Jane, Wiser and the proposed process in this paper. This table considers the currently available tools, not the techniques that are behind them. For example, Wiser can be used on different data sources, but in the currently available version, it is based only on University of Pisa publications. *Publish or Perish* is not an expert finding tool, it is an effective citation analysis tool that can be used for expert finding purposes. We recommend to import output data of Publish or Perish in BI tools for expert finding purposes.

Name*	Data source	Level of required knowledge	Capability for defining criteria by user	Visualization capability	Expert ranking
Jane	PubMed	No special knowledge, easy to use	Limited criteria can be defined based on advanced search option in the tool UI	No	Yes, automatically
Wiser	University de Pisa publications	No special knowledge, easy to use	There is no option for defining criteria in wiser UI	Yes	Yes, automatically
Publish or Perish	Web of science, Scopus, Crossref Google Scholar, and Microsoft Academic Search. It is also possible to import external data	primarily knowledge about scientific bases and citation analysis is necessary	User can define some criteria	No	It is possible to rank expert based on output values. For example, sorting based on h-index
Proposed process	Publications data from different sources such as Scopus, or Google Scholar	Primarily knowledge about data, scientific databases and data tools necessary	User can define different criteria as there is data to support such criteria	Yes, by using BI tools visualization elements	Yes, manually by using AHP and automatically by defining in BI tools

This process can be compared with two main expert finding approaches: manual expert finding by searching in scientific databases and proposed technical methods in the literature. Researchers can use scientific databases such as Google Scholar or Scopus to search for keywords and manually inspect search result to find experts. The process in this paper has other advantages including:

- In the manual inspection of result, researchers cannot consider all results and are limited in the publications that they can analyze in terms of time and effort.
- Researchers cannot execute an advance query on search result without utilizing BI tools without advanced technical knowledge.
- When data are collected from other sources, such as organizational publications or internal repositories, it is not possible to use Scopus or scientific databases to import data for analysis.
- When data come from internal repositories, they may be in different topics and domains, thus, manual inspection of such data may require significant time and effort to assess.

In comparison with proposed technical methods in the literature for expert finding, this process is easier in terms of implementation for researchers who have BI knowledge but do not have advance technical knowledge. If technical methods are the tool implemented and are publicly accessible for all researchers, they can be compared in terms of capabilities and advantages with BI tools. For that purpose, a comparison between Jane, Wiser and the process in this paper is shown in Table 5.

The process in this study may be limited to cases where data is related to the potential experts' publications. Future research can focus on using BI tools to find experts based on data gathered from social networks or community question answering websites. In this paper we only focus on the usefulness and level of required technical knowledge to evaluate this process with the proposed methods in the literature. The main goal of this study is to propose a simpler expert finding process, which provides acceptable results based on analyzing publications, not providing a comprehensive expert finding method. The contribution of this research is a discussion on a process for finding experts by using BI tools. This paper does not propose a new tool or method, but it introduces the capability of existing BI tools for finding potential experts.

4. CONCLUSION

Given that the existing expert selection methods are usually impractical for researchers without deep technical knowledge, an expert selection process is discussed here for individuals who are familiar with BI tools. Taking advantage of BI tools, the process was found to have a large potential for expert finding. The process will be helpful in research that aims to gather data from expert participants. Here, we may need the opinions of experts and finding these experts is key.

The process in this paper requires a certain level of technical knowledge, because the method for expert finding is based on computers, which are technical in nature. The primarily knowledge about data, scientific databases and data tools is necessary for individuals who aim to use BI tools for expert finding. However, such knowledge can be obtained by participating in a workshop or reading relevant books and tutorials. This process is simpler, when we are aware of BI tools that support different options to simplify tasks, such as providing drag and drop options (Smuts, Scholtz, & Calitz, 2015). In addition, there are efforts for providing self-service BI tools which individuals can use with less technical knowledge (Imhoff & White, 2011). However, utilizing expert knowledge of programming helps researchers to collect more complete data and execute more complex queries. Also, for advanced data analysis, the knowledge of programming may be essential. Researchers, by improving their skills, could gain more benefit from this process. BI tools have the potential for data visualization and analysis, but related skills are required for such capabilities be reachable. In this paper, BI tools have been used to find an early list of potential experts from the data, then AHP helps to manually distinguish them and produce a final list of experts. Based on available data, a primary filtering of the list of many experts is done through BI tools, then by using AHP, a final list of experts is identified manually. So, queries in the BI tool may be simple, for example finding experts who have a total of more than 1000 citations. Such queries will make a limited list of potential experts, which is usable in AHP. The threshold and criteria for early filtering of experts using BI tools can be defined by consulting with experts. All thresholds in the presented cases in this paper are examples. In the actual expert finding process, consulting with experts to identify threshold and selection criteria based on available data for early filtering of experts is required. This process helps researchers to find experts for their work, even they are not experts in BI tools. However more knowledge and skills are needed for BI tools, to make them more successful in finding suitable experts.

ACKNOWLEDGEMENTS

It is our pleasure to thank DBxtra company for their support and providing a trial license of their tool for our work. We also appreciated assistance of researchers who helped us to understand their methods or answered our questions regard to this research. They are: Professor Robert Davison, City University of Hong Kong; Professor Dr. Muhammad Tanvir Afzal, Department of Computer Science, Capital University of Science and Technology, Islamabad, Pakistan; Professor Dr. h.c. mult. Hermann Maurer, Computer Science, Graz University of Technology, Graz/ Austria; Dr. Muhammad Naeem, researcher in mobility, VeDeCoM, Versailles France; Prof. Steven Gordon, Technology, Operations and Information Management Division, Babson College, Babson Park, MA; and other researchers and companies who helped us.

5. REFERENCES

Afzal, M. T., Kulathuramaiyer, N., & Maurer, H. (2008). Expertise finding for an electronic journal. *proceedings of I-Know (Graz, Austria*, 436-440.

Afzal, M. T., & Maurer, H. A. (2011). Expertise Recommender System for Scientific Community. *J. UCS, 17*(11), 1529-1549.

Beatty, S. (2015). Analysis tools-Analyze thousands of search results in less than a minute Retrieved from https://blog.scopus.com/topics/analysis-tools

Benner, P., Tanner, C., & Chesla, C. (1992). From beginner to expert: gaining a differentiated clinical world in critical care nursing. *ANS Adv Nurs Sci, 14*(3), 13-28.

Bernardino, J., & Tereso, M. (2013). *Business Intelligence Tools*, Dordrecht.

Boeva, V., Angelova, M., & Tsiporkova, E. (2017). *Data-driven Techniques for Expert Finding.*

Paper presented at the International Conference on Agents and Artificial Intelligence.

Bozzon, A., Brambilla, M., Ceri, S., Silvestri, M., & Vesci, G. (2013). *Choosing the right crowd: expert finding in social networks*. Paper presented at the Proceedings of the 16th International Conference on Extending Database Technology, Genoa, Italy.

Cameron, D. H. L., Aleman-Meza, B., Decker, S., & Arpinar, I. B. (2007). *SEMEF: A taxonomy-based discovery of experts, expertise and collaboration networks.* University of Georgia.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Commun. ACM, 54*(8), 88-98. doi:10.1145/1978542.1978562

DBxtra. (2018). DBxtra Online Documentation. Retrieved from www.dbxtra.com/documentation/

Demartini, G. (2007). *Finding experts using wikipedia.* Paper presented at the Proceedings of the 2nd International Conference on Finding Experts on the Web with Semantics-Volume 290.

Deng, H., King, I., & Lyu, M. R. (2008, 15-19 Dec. 2008). *Formal Models for Expert Finding on DBLP Bibliography Data.* Paper presented at the 2008 Eighth IEEE International Conference on Data Mining.

Flick, U. (2008). *Designing Qualitative Research*: SAGE Publications.

Girvan, C., & Savage, T. (2013). Guidelines for Conducting Text Based Interviews in Virtual Worlds. In M. Childs & A. Peachey (Eds.), *Understanding Learning in Virtual Worlds* (pp. 21-39). London: Springer London.

Gretsch, S., Mandl, H., & Hense, J. (2011). *The Difficulty of Finding Experts-Implementation Process of Corporate Yellow Pages.* Paper presented at the KMIS.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569-16572. doi:10.1073/pnas.0507655102

Imhoff, C., & White, C. (2011). Self-service business intelligence: Empowering users to generate insights. *TDWI best practices report*, 40.

Isaak, D. (2016). PubMed2XL (version 2.01). *Journal of the Medical Library Association: JMLA, 104*(1), 92.

Kao, W.-C., Liu, D.-R., & Wang, S.-W. (2010). *Expert finding in question-answering websites: a novel hybrid approach*. Paper presented at the Proceedings of the 2010 ACM Symposium on Applied Computing, Sierre, Switzerland.

Kardan, A., Omidvar, A., & Farahmandnia, F. (2011, 17-19 May 2011). *Expert finding on social network with link analysis approach.* Paper presented at the 2011 19th Iranian Conference on Electrical Engineering.

Lin, S., Hong, W., Wang, D., & Li, T. (2017). A survey on expert finding techniques. *Journal of Intelligent Information Systems, 49*(2), 255-279. doi:10.1007/s10844-016-0440-5

Louise Barriball, K., & While, A. (1994). Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing, 19*(2), 328-335. doi:doi:10.1111/j.1365-2648.1994.tb01088.x

Markham, A. N. (2004). Internet communication as a tool for qualitative research. *Qualitative research: Theory, method and practice, 2*, 95-124.

Marshall, M. N. (1996). Sampling for qualitative research. *Family practice, 13*(6), 522-526.

Matsuo, Y., & Ishizuka, M. (2004). Keyword Extraction From A Single Document Using Word Co-Occurrence Statistical Information. *International Journal on Artificial Intelligence Tools, 13*(01), 157-169. doi:10.1142/s0218213004001466

Merrouni, Z. A., Frikh, B., & Ouhbi, B. (2016, 24-26 Oct. 2016). *Automatic keyphrase extraction: An overview of the state of the art.* Paper presented at the 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt).

Naeem, M., Khan, M. B., & Afzal, M. T. (2013). Expert Discovery: A web mining approach. *Journal of AI and Data Mining, 1*(1), 35-47. doi:10.22044/jadm.2013.116

Negash, S. (2004). Business intelligence. *Communications of the association for information systems, 13*(1), 177-195.

OpenRefine. (2018). Retrieved from http://openrefine.org/

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful Sampling for Qualitative

Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533-544. doi:10.1007/s10488-013-0528-y

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Quatrini Carvalho Passos Guimarães, H. C., Pena, S. B., Lopes, J. d. L., Lopes, C. T., & Bottura Leite de Barros, A. L. (2016). Experts for Validation Studies in Nursing: New Proposal and Selection Criteria. *International Journal of Nursing Knowledge, 27*(3), 130-135. doi:doi:10.1111/2047-3095.12089

Riahi, F., Zolaktaf, Z., Shafiei, M., & Milios, E. (2012). *Finding expert users in community question answering*. Paper presented at the Proceedings of the 21st International Conference on World Wide Web, Lyon, France.

Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic keyword extraction from individual documents. *Text Mining: Applications and Theory*, 1-20.

Ru, Z., Xu, W., & Guo, J. (2007). *An Expert Experience Probabilistic Model for Enterprise Expert Finding*. Paper presented at the Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007).

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology, 15*(3), 234-281.

Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research, 48*(1), 9-26.

Saaty, T. L. (2013). Analytic Hierarchy Process. In S. I. Gass & M. C. Fu (Eds.), *Encyclopedia of Operations Research and Management Science* (pp. 52-64). Boston, MA: Springer US.

Serdyukov, P., & Hiemstra, D. (2008). *Modeling documents as mixtures of persons for expert finding*. Paper presented at the European Conference on Information Retrieval.

Sherman, R. (2014). *Business Intelligence Guidebook: From Data Integration to Analytics*. Newnes.

Smuts, M., Scholtz, B., & Calitz, A. (2015). *Design guidelines for business intelligence tools for novice users*. Paper presented at the Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists.

SuperDecsion. (2018). Retrieved from http://www.superdecisions.com/

Verborgh, R., & De Wilde, M. (2013). *Using OpenRefine*: Packt Publishing Ltd.

Wang, G. A., Jiao, J., Abrahams, A. S., Fan, W., & Zhang, Z. (2013). ExpertRank: A topic-aware expert finding algorithm for online knowledge communities. *Decision support systems, 54*(3), 1442-1451.

Yang, K.-H., Chen, C.-Y., Lee, H.-M., & Ho, J.-M. (2008). *EFS: Expert finding system based on Wikipedia link pattern analysis*. Paper presented at the 2008 IEEE International Conference on Systems, Man and Cybernetics.

Zhang, J., Tang, J., & Li, J. (2007). *Expert Finding in a Social Network*, Berlin, Heidelberg.

Zhang, J., Tang, J., Liu, L., & Li, J. (2008). *A Mixture Model for Expert Finding*, Berlin, Heidelberg.

Zhao, Z., Zhang, L., He, X., & Ng, W. (2014). Expert finding for question answering via graph regularized matrix completion. *IEEE Transactions on Knowledge and Data Engineering, 27*(4), 993-1004.