St. Petersburg University

Graduate School of Management

Master in Management Program

DEVELOPMENT OF A METHOD  
FOR EXPERTISE LOCATION AT UNIVERSITIES

Master’s Thesis

by the 2nd year student

Concentration —

International Business

Aleksandr Andreev

Research advisor:

Associate Professor

Dmitry Kudryavtsev

St. Petersburg

2020

## 

## 

# АННОТАЦИЯ

|  |  |
| --- | --- |
| Автор | Андреев Александр Алексеевич |
| Название ВКР | Разработка метода идентификации экспертов в университетах |
| Образовательная программа | Высшая школа менеджмента |
| Направление подготовки | Международный бизнес |
| Год | 2020 |
| Научный руководитель | к.т.н., доцент, Кудрявцев Дмитрий Вячеславович |
| Описание цели, задач и основных результатов | Цель данной работы заключается в разработке метода идентификации экспертов в университетах, позволяющего решать широкий спектр проблем за счет тонкой настройки в соответствии с запросом и использования нескольких релевантных источников. Исследование носит качественный характер и использует методологию Design Science Research (DSR).  Для достижения этой цели были поставлены следующие задачи:  • определить основные процессы в университетах, которые могут быть улучшены путем внедрения метода идентификации экспертов;  • перечислить и проанализировать известные методы и приемы процесса идентификации экспертов и оценить их применимость к процессам, выявленным на предыдущем этапе;  • разработать метод идентификации экспертов, применимый к любому высшему учебному заведению на основе существующих методов и приемов;  • продемонстрировать этот метод на экспертах определенного университета и подтвердить корректность и точность результатов.  В рамках исследования был разработан метод идентификации экспертов, который позволяет более гибко настраивать процесс определения местоположения экспертизы, решать широкий спектр проблем, имеющих отношение к академической среде, повышать эффективность значительной доли внутренних процессов, поддерживать принятие стратегических решений и управление человеческими ресурсами посредством отображения общей экспертизы организации. Предложенный метод учитывает особенности академической среды и был применен для создания профилей экспертизы нескольких экспертов Высшей школы менеджмента СПбГУ. |
| Ключевые слова | Управление знаниями, поиск экспертов, высшие учебные заведения |

# ABSTRACT

|  |  |
| --- | --- |
| Master Student's Name | Aleksandr A. Andreev |
| Master Thesis Title | Development of a Method for Expertise Location at Universities |
| Educational Program | Graduate School of Management |
| Main field of study | International Business |
| Year | 2020 |
| Academic Advisor’s Name | Candidate of Engineering Sciences, Associate Professor, Dmitry V. Kudryavtsev |
| Description of the goal, tasks and main results | The goal of this master thesis is to develop an expertise location method for universities allowing to solve various user problems through customization and using multiple relevant sources. This research is of qualitative nature and uses Design Science Research (DSR) methodology.  To achieve this goal, following objectives were set:  • To identify main processes in the universities that can be enhanced by introducing expertise location;  • To list and analyze known methods and techniques of expertise location process and estimate their applicability to the processes in the previous step;  • To develop a method for expertise location, applicable for any academic environment, considering existing methods and techniques;  • To demonstrate this method on experts from the same university and verify adequacy and precision of results.  As a result, an expertise location method was developed. Based on existing tools and techniques, it allows for more customizable expertise location process, solving a wide range of needs relevant for academic environment, enhancing a major share of internal processes, supporting strategic decision making and human resource management through depicting overall organization expertise. The suggested method considers features of academic environment and was demonstrated by being applied to create expertise profiles of several experts of Graduate School of Management of Saint-Petersburg State University. |
| Keywords | Knowledge management, expertise location, universities |

# TABLE OF CONTENTS

[TABLE OF CONTENTS 2](#_Toc42101764)

[INTRODUCTION 4](#_Toc42101765)

[CHAPTER 1. THEORETICAL BACKGROUND OF EXPERTISE LOCATION PROCESS 6](#_Toc42101766)

[Knowledge mapping 6](#_Toc42101767)

[Expertise location systems 15](#_Toc42101768)

[CHAPTER 2. RESEARCH GAP AND METHODOLOGY 21](#_Toc42101769)

[Research gap 21](#_Toc42101770)

[Research goals and objectives 22](#_Toc42101771)

[Design science research methodology 22](#_Toc42101772)

[Data collection methods 24](#_Toc42101773)

[CHAPTER 3. METHOD DESIGN AND DEVELOPMENT 25](#_Toc42101774)

[Purposes of expertise location and valid sources of expertise 25](#_Toc42101775)

[Requirements for the method of expertise location 26](#_Toc42101776)

[General concepts of the method of expertise location 28](#_Toc42101777)

[Method for expertise location 30](#_Toc42101778)

[CHAPTER 4. METHOD DEMONSTRATION 34](#_Toc42101779)

[Description of organization and environment 34](#_Toc42101780)

[Identifying relevant goals of expertise search 35](#_Toc42101781)

[Gathering primary data 36](#_Toc42101782)

[Data processing and intermediate results 39](#_Toc42101783)

[Query-based final stage of data processing and visualization 42](#_Toc42101784)

[DISCUSSIONS 45](#_Toc42101785)

[Benefits and limitations of the method 45](#_Toc42101786)

[Suggestions for further research 46](#_Toc42101787)

[CONCLUSION 47](#_Toc42101788)

[REFERENCES 48](#_Toc42101789)

[Appendix 1 54](#_Toc42101790)

[Appendix 2 56](#_Toc42101791)

[Appendix 3 58](#_Toc42101792)

[Appendix 4. 59](#_Toc42101793)

# INTRODUCTION

In current rapidly changing environment with growing complexity of all processes and products it is becoming crucial to manage knowledge assets both for individuals in order to be able to solve challenging problems and increase personal effectiveness and for organizations to gain competitive advantage, assess and mitigate risks caused by concentration of knowledge within several experts by capturing this knowledge and making it easily accessible throughout the company, to determine knowledge gaps to further increase overall organization effectiveness by educating employees to internalize potentially important knowledge, to understand and describe organization core competences and capabilities.

. Knowledge management as a field of study has existed for more than 30 years. The concept of knowledge primarily was related mostly to the academic field, but now it has become a crucial element of organizational life (Kapur, 2020).

Knowledge management aims to develop tools, practices and frameworks to find, extract and effectively manage knowledge, which is also developing rapidly and gaining more attention lately. One of important problems that knowledge management tries to solve is that a great deal of knowledge is tacit and either is not formally described or is impossible to be extracted, being accessible only through reaching for a suitable expert.

Knowledge mapping is a powerful method enabling company to connect experts, access knowledge in time, identify knowledge assets and flow, identify existing knowledge resources and knowledge gaps. It also can be utilized to show the important stages to build up a specific capability (Faisal et al., 2019).

Main tools that are most widely used in knowledge mapping require participation of several experts, which takes time and lays restrictions upon the possible frequency of updating the map. At the same time keeping knowledge map as relevant as possible is crucial, since relying on outdated map can lead to negative consequences such as impossibility to get access to knowledge or wrong decisions in terms of managing HR (e.g. developing outdated competences or irrelevant skills), which undermine the whole concept.

Another problem is that proper knowledge map has to be perfectly balanced between simplicity and sufficiency of detail, so that it could be easily grasped at organization level and at the same time useful at operation level. In order to achieve this, we have to design the map for the particular organization and for the particular purposes. We suppose that there are ways that allow educational organizations to identify experts in specific fields by analyzing their publications, participation in conferences or some other activities that can be automatically gathered and processed to create and update at least some types of knowledge maps without need to conduct personal interviews consuming significant amount of time and effort.

But in order to draw a map, showing where knowledge is created, located and describing how it flows throughout organization we have to identify experts first. That means we need to assess expertise of each potential expert in each domain of knowledge relevant for a specific organization.

Moreover, expertise location system might be helpful even without further integration of its results into knowledge maps. People often prefer to consult with an expert when they face a problem they do not know how to solve so providing them with a useful tool to do that is a good application of plain expertise location system.

# CHAPTER 1. THEORETICAL BACKGROUND OF EXPERTISE LOCATION PROCESS

## Knowledge mapping

The concept of knowledge maps seems to many to be not quite understandable due to the lack of widespread adoption of generally accepted concepts. At the same time, knowledge maps are already deeply integrated in our everyday life, but due to the lack of formalization in writing, they are not perceived by us as the phenomenon which might be subject of research. Indeed, we all have the knowledge who is the right person to advice or consult in a particular situation, thus, we understand who potentially has the necessary knowledge in a particular area. However, if we switch subject of this case from an individual to an organization, we will see that there are too many areas of knowledge and experts, and moreover, in large organizations, employees often simply do not know who to contact, because they did not personally contact experts in this field. This leads to a long chain of request redirects. In this case, knowledge maps are becoming a crucial tool that allows to document every grain of knowledge inside the object of mapping, in our case, the organization, and helps any user of the map to find any necessary existing information.

Nevertheless, there is no single universally accepted definition of knowledge maps, so many researchers in their works rely on their own understanding of the concept of knowledge maps. So let us review some of the definitions and agree which one we will use in this review (table 1).

Table 1. Definitions of knowledge map

|  |  |
| --- | --- |
| **Definitions** | **References** |
| “A knowledge map serves as continuously evolving organizational memory, capturing and integrating strategic explicit knowledge within an organization and between the organization and its environment” | Wexler (2001) |
| “A knowledge map is a navigation aid for discovering the sources of explicit and tacit knowledge by illustrating how knowledge flows through the organization” | Chan & Liebowitz (2006) |
| “A knowledge map portrays ‘the sources, flows, constraints and sinks of knowledge’ within the organization” | Liebowitz (2005) |
| “A knowledge map is about making the knowledge that is available within an organization transparent and providing insight into its qualities” | Driessen, Huijsen and Grootveld (2007) |
| “A knowledge map is about discovering knowledge, tracing its flow, mapping its existence and its changes, and identifying where it is most needed” | Lee & Fink (2013) |

For the academic domain and for expertise location aspect the most appropriate is the definition, offered by Driessen, Huijsen and Grootveld (2007), since knowledge flow and application are outside of this thesis scope, and this definition mentions all key points, crucial for this research.

Based on existing research, it is safe to say that knowledge management is a large interdisciplinary field, which can be recapitulated in the phrase ‘know, show, grow!’. Know is tacit “knowledge in an individual’s head”; otherwise knowledge that is documented and written down means to show (explicit) and grow is considered as a collaboration toward innovation that stimulates new knowledge. (Balaid et al., 2016) Among five segments of knowledge management, as described by Duffy (2000), knowledge maps refer to the segment of user interface and are confirmed to have the potential to be a primary component of knowledge management.

Tiwana refers to knowledge mapping as the tool for getting an overview of the strategic position within organizations and industries (Tiwana, 1999). Chung, Chen and Nunamaker (2003) also agreed with Tiwana in that a knowledge map should be used on the top level of organizational issues, such as strategy, vision and mission. Further, the researchers (Hansen & Kautz, 2004) stated that knowledge maps are a crucial prerequisite for effective knowledge management. Furthermore, the researchers (Hellström & Husted, 2004) share the same vision that knowledge maps can be used to effectively create, capture, codify, transfer, retain and use knowledge in organizations and projects.

From a different point of view, Earl (2001) classified the strategies of knowledge into a number of schools: “economic”, “technocratic” and “behavioral”. The cartographic school, which is the part of technocratic school, clearly connected with the concept of a knowledge map since it focuses particularly on maps. Engineering and systems schools which are also included to technocratic school also fit the idea of knowledge mapping since they are based on information technologies that can help all employees throughout the organization in their activities. To sum up, a knowledge map is considered as a key solution for successful knowledge management in an organization since it gives users a clear understanding of the knowledge location and flow throughout organization.

Knowledge maps technique is a powerful tool of knowledge management that allows to visualize how the knowledge is created within the company, how it is applied, and how it flows throughout organization, as well as how it is stored and accessed.

Knowledge maps lately are recognized as a crucial subfield of knowledge management and draw attention from researchers. They describe what useful information exists within an organization and organize this description in a way that enables user to have a grasp of how and where to find it (Eppler & Simon, 2008). Using knowledge maps might help to achieve a number of goals. Some organizations use this tool for strategic planning purposes, others use it to enable knowledge flow throughout organization, leading to “empowerment of the organization” which “involves providing access to existing information and expertise” (Hellström & Husted, 2004).

Sometimes knowledge maps are used to depict the views of participants and their mutual relations to other views, as well as illustrate the dependencies of learning paths and serve as the basis for the implementation of knowledge management programs (Dang et al., 2011).

In all these options for using knowledge maps, they solve the problem of the most effective visualization of existing knowledge in order to use it to achieve a specific goal, especially when organizations that recognize the importance of effective management of knowledge assets would like to know how, when and where to access them. Knowledge map presents a snapshot of where an organisation is at any given time which could be further used to elaborate on core competitive advantages as well as to identify and eliminate key weaknesses (Wexler, 2001).

Ali Balaid et al. (2016) in their literature review distinguish 7 directions of currently conducted research of knowledge maps, which are:

* benefits and utilizations of knowledge maps;
* knowledge maps systems and tools development;
* knowledge maps frameworks and models;
* methodologies of building knowledge maps;
* knowledge maps classification;
* knowledge maps adoption;
* general reviews of knowledge maps.

Using knowledge mapping within organization can generate multiple benefits. Wexler (2001), classified the benefits of a knowledge map into four major clusters:

* knowledge;
* economic;
* structural;
* organizational/cultural.

Yasin and Egbu (2011), grouped the benefits of a knowledge map into:

* cost saving;
* process improvements;
* knowledge improvement;
* value improvement;
* user satisfaction.

Eppler (2001), stated that the most obvious benefit and the major goal of a knowledge map is to show people in an organisation where to look for required information and who possesses such expertise. Moreover, a knowledge map can increase transparency and reduce complexity within organizations as well as eliminate overlapping and repetitive activities (Wexler, 2001; Yasin & Egbu, 2011).

According to Rao, Mansingh, and Osei-Bryson (2012), a knowledge map can support management in addressing key organizational issues, identifying knowledge gaps, establishing clear links between existing knowledge assets, thus enhancing the decision-making process.

In summary, the literature suggests a broad range of knowledge maps benefits; however, the findings provide a clear account of the key benefits, as reported in the literature, although they may not be complete or mutually exclusive due to the approach applied. Thus, to overcome the limitations future research is encouraged.

Most researchers name the following benefits of knowledge maps: connecting experts, accessing knowledge in time, identifying knowledge assets, identifying knowledge flow, identifying existing knowledge resources, organizational restructure, identifying knowledge gaps, team building, identifying untapped knowledge.

There are some tools and techniques that are most commonly used, which are classified by Ali Balaid et al. in the Table 2:

Table 2. Classification of knowledge-mapping tools and techniques

|  |  |  |
| --- | --- | --- |
| **Techniques/tools** | **Description Application** | **References** |
| Concept map | The structure of this technique is “node, link” structure where nodes represent concepts and likes illustrate the concepts relationships. It is the best technique for visualizing the relationships among diverse concepts. | Can be applied to express knowledge of a particular person about certain topic in a specific domain, to clarify misconceptions and in some cases for brainstorming and generating new ideas. |
| Mind map/idea map | This technique/tool is an image centered drawing that used for showing the connections among portions of information. | Can be applied as a technique for note-taking, reflect or capture the process in brain, problem solving and training etc. |
| Concept circle diagram | Concept circle diagram techniques/tools, are labelled circles, which possibly inclusive, exclusive, and/or overlapping to present concepts relationships. | This tool applied to group ideas into categories, showing the existing relationship between themes, departments as well organizations. |
| Semantic map | This technique/tool is more similar to concept map and mind map that mentioned above | It is applied for machine translations and AI (artificial intelligence). |
| Cognitive map | This maps are used for mapping the person’s thoughts about specific subject or situation. As well it is used for enhancing learning and recall information. | Applied in referring to models of mental that people used for perceiving, contextualizing, simplifying, |
| Process map | This technique represents blocks of tasks or activities in sequenced and logical way to achieve particular goal. | Applied to identify tasks sequences steps to achieve a particular goal. |
| Conceptual map | It is a graphical means used to arrange concepts and conceptual relations in understandable way. | Utilized as a communication language between computer systems and peoples |
| Knowledge flow map | This technique/tool is a high-level knowledge-model in obvious graphical form. | Applied to map the flow of knowledge and measure the relationship between  teams, groups, departments and firms |
| Causal map | This tool used link graphs and directed node to show a set of causal relationships within the system. | Used to explore group’s or individual’s beliefs so as to establish cause & effect relationships. |
| Social mess map | Used to summarize the understanding of a particular group about a given problem. | Applied to summarize team or group understanding |
| Ontology | llustrating concepts and the existing relationships for a community of agents. The main types are task-oriented, domain-oriented, and generic. | Used as a tool to find all information available in a given topic. |
| Petri net | This technique is abstract formal model of knowledge flow, composed of transitions, directed arcs, and places. | Applied in search for simple, natural, powerful methods for analyzing and showing information flow in systems |
| Cluster Vee diagram | This map form of V-shaped. Used as a road map to show the route of knowledge from the past to future. | Helping students in science field to make the essential elements unambiguous to build a scientific knowledge |
| Thesauri | This map is a set of concepts characterized by synonymous, horizontal, hierarchical, and other relations. | Often applied in retrieval system. |
| Visual thinking network | This is a knowledge’s representation strategy that encourages the users to combine multiple ways of thinking that put in the picture concept formation. | Used to organize, revise and represent user’s knowledge meaning-making by linking and grouping pictorial visualizations and symbolic into a coherent whole. |
| Topic map | This technique helps users to find the direct path to a specific information within the concept (is more or less a book indexes electronic version). | Used to manage large quantities of unorganized information. |
| Perceptual map | This technique often used to take the complicated results from research surveys and showing them on obvious and informative map. | Usually used by marketers to display customers or potential customers perceptions. |

There are several classifications used in industry and academic research. Logan and Caldwell (2000), distinguish three types of knowledge maps:

* conceptual knowledge maps, which depict relationships between knowledge sources within the organization;
* competency knowledge maps, which shows relationships between people and knowledge;
* process knowledge maps, which.describes procedures and schedules within organization.

Gilbert and Raub (2000) used a different classification, and divided knowledge maps into:

* source knowledge maps are widely used to present information about internal (in some cases even external) experts that possess information, crucial for specific task;
* assets knowledge maps strive to document, codify and organize all knowledge assets within organization in order to make them visible and easily obtainable when needed;
* topography knowledge maps define knowledge, including information such as how much, what and whom.

Eppler (2001) reworked this classification by replacing topography knowledge maps with three more specific types of knowledge maps:

* knowledge application map connects location of particular knowledge on one side and specific business situation or a certain process stage on other;
* knowledge development map shows necessary steps to develop desired competence;
* knowledge structure map contains structure of the knowledge domain, including relationships between its subdomains.

Source and assets knowledge maps that together with three aforementioned types comprise for the classification, are similar to those in Gilbert and Raub (2000).

Huff (2002) concentrates on cognitive maps and refers to the following types of maps:

* schematic maps of cognitive structures;
* text and language analysis maps;
* network maps;
* classification maps;
* conclusive maps.

Vestal (2005) referred to three major groups of knowledge maps:

* enterprise knowledge maps based on expertise overviews and strategic overviews of knowledge maps;
* process explicit knowledge maps, which include expertise gaps, document-explicit and job role based knowledge maps;
* cross-functional knowledge maps, composed of functional, technical and tacit knowledge maps.

Eppler and Simon (2008) proposed to classify knowledge maps by their intended purpose, content, application level, creation method, or by their graphic form.

Prior research in a knowledge map area, as discussed by Kim et al. (2003) and Wexler (2001), highlighted importance of selecting the best methodology for building a knowledge map for each particular case and purpose. Balaid et al. (2016) in their search for methodologies in the studies that focus on knowledge maps (the primary set of 132 studies) identified seven different types of methodology for developing knowledge maps from seven different authors, as summarized below. Each methodology was not necessarily referred to as a knowledge map methodology but was considered to operate in a similar way in the context of this report.

The first methodology from (Bargent, 2002) has eleven steps and is based on the strategy of software development lifecycle stages by using the Lotus Discovery Server (LDS) software, then identifying requirements and finally testing the knowledge map. This method has strong validation stage, which is due to usability tests used to estimate the effectiveness of the knowledge map as well as to ensure that all information is documented and classified before analysis. The disadvantage is that the methodology requires the use of LDS, which makes it difficult for employees with little or no experience in software application to contribute.

Kim et al. (2003) suggest six stages methodology, which employs the use of ontology, which is a suitable tool for formally representing knowledge based on a conceptualization in relation to the objects and concepts that exist within the specific domain and the relationships between them. This method was tested in the manufacturing industry.

Lecocq (2006) created methodology, consisting of four phases, which are: planning, collecting, mapping and validating, which, in turn, could be broken down into eleven steps. This method is quite consistent due to the use of ontology as well, but is more detailed than previous ones.

Mansingh, Osei-Bryson, and Hirata (2009) developed three steps methodology, which assumes ontology creation, processes identification and bringing their examples. All existing types of knowledge are then presented as the final knowledge map. The method was tested in a healthcare organisation.

Pei and Wang (2009) proposed the seven steps methodology: project team formation, knowledge assets analysis, knowledge domain identification, structure and relationships definition, development tools selection and evaluation, knowledge classification confirmation and knowledge map constant updating. The method was tested in a matrix organization.

Liu, Li, and Lv (2009) proposed the three-level methodology, which results in the knowledge map for organization as a multilevel system (individual, domain and interdomain levels). The methodology was developed by combining these elements to construct an organization knowledge map and then applied to a real-world case, which confirmed methodology fit.

The last methodology was developed by Zhao, Miao, and Guan (2013), which was designed for complex product development and consists of seven steps:

* building context awareness;
* formulating purpose along with related information;
* understanding the product structure, transforming the context information to specific knowledge requirements;
* defining most important characteristics of the knowledge and designing a knowledge retrieval process: formalization, location and classification;
* selecting the knowledge point;
* determining the most effective visualization;
* drawing a knowledge map from different aspects.

This methodology can help to overcome the difficulty of creating a full set of knowledge maps for each aspect of knowledge mapping concept while meeting the requirements.

From the analyzed researches, it is clear that knowledge maps practical usage is at the early stage of its development and still has challenges to overcome. According to Renukappa and Egbu (2004) and Dang, Zhang, Chen et al. (2011) the most relevant of such challenges are the mapping of tacit and dynamic knowledge, the mapping of interdomain knowledge, organizational culture and representation of knowledge.

Generally, knowledge management is especially challenging due to its dynamic nature: some knowledge may get outdated and replaced with more relevant, while other depreciates over time, and some may simply be lost due to some reasons. This calls for a multifunctional approach and dynamic knowledge mapping (Renukappa & Egbu, 2004). Another significant problem for knowledge mapping is linked with efficient tacit knowledge extraction and codification into computer processable data. Miscommunication in maps, when the users and creators of the knowledge map do not have the opportunity to personally communicate or disagree on what the purposes of the knowledge map are, is also a huge problem. Since a knowledge map is abstract in its nature it should reflect reality in a way that must be understood and shared by everyone who creates or uses it (Eppler & Simon, 2008). Knowledge mapping of cross-boundary knowledge is also a crucial task. Thus, organizations must develop competences to precisely represent complex knowledge irrespectively of both internal and external boundaries and between different domains (Quintas, 2002). In addition, organizational leadership and culture should also be included into the basis of knowledge mapping for better results. Not motivated or inefficient management often leads to lack of money, time, and talent resources, which most probably will restrict and undermine the whole process of knowledge mapping.

Eppler (2004) mentions following challenges as the most important: organizing the knowledge with the relevant references in such framework that everyone can access, information overload problem, simplifying complex structures to the degree that it can be visualized in easy to grasp format without losing important details, maps becoming outdated over time, and the risks that can emerge if the map is seen by unauthorized users, such as competitors or recruiting agencies looking for an expert.

Vestal (2005) refers to the following key challenges that could possibly arise along knowledge maps implementation process: lack of understanding of the flow of knowledge within the organization, the wrong selection of the team that creates knowledge maps, ‘knowledge is power’ motto increasing knowledge “stickiness” and blocking its transfer from experts through the whole organization, and the failure to understand the business process.

Some of these obstacles are personal and cannot be managed directly, yet management can still somehow control them by applying pull and push strategies to help workers overcome their concerns and pull them towards utilizing knowledge maps (Balaid et al., 2016).

## Expertise location systems

In order to build knowledge maps and utilize their benefits it is required to extract all knowledge that exists within organization or relevant for knowledge map that is being developed. Expertise is an important knowledge asset which is often tacit and so is hard to extract and codify, which leads to the necessity to apply special methods to do that. Expertise can be distributed when people communicate and interact. Users need to consult an expert to get insights on how to solve their problems, which lie in specific area. Though usually there is a lot of information available that can help in this case, mostly people still are more comfortable with the advice of an expert. Thus, conference participants search for paper reviewers, and students need appropriate research supervisors (Balog et al., 2012). The term ‘expertise’ in psychology is defined as the human cognitive skills that can be developed by conducting a task frequently (Anderson, 1999). Generally, people who possess expertise in a particular domain are called experts. However, there is no widely accepted definition of the term ‘expert’. Weiss and Shanteau (2003) defined an expert as an individual who has skills in his domain of expertise and can evaluate a particular domain topic. Furthermore, Lin et al (2017) identified an expert as an individual who has a great deal of knowledge and displays comprehensive skills in a specific area. Naeem, Khan and Afzal (2013) identified an expert as a person who has the best experience about a particular topic in a particular field of study, based on a collection of various factors.

The expertise seekers are those people who are looking for knowledge for particular purposes such as problem-solving and collaboration. They require all the necessary information about experts who can help them with relevant expertise. It is ineffective to manually gather data about experts, especially in highly fragmented and large-scale organisations (Mangaravite et al., 2016, Yimam-Seid and Kobsa, 2003). Therefore, Information Retrieval (IR) techniques are used to automate the selection of the suitable experts. Such techniques applied for this purpose are called expert finding systems (also referred to as expert locating or expertise retrieval systems). Expert finding systems identify candidate experts and rank them by their expertise in a specific domain. The candidate’s expertise is extracted from their expertise evidences (expertise sources) such as publications, reports, projects, and social network (Stankovic et al., 2010, Balog, Azzopardi and De Rijke, 2009). The expertise retrieval system is an effective tool for both individuals and organisations, as it enables them to easily find appropriate experts, based on the expert’s actual knowledge (Ehrlich, 2003). Early expert finding systems were based on particular expertise sources. Thus, Mcdonald and Ackerman (2000) and Mcdonald (2001) studied locating experts in the technical and support departments of a software corporation. Mockus and Herbsleb (2002) created a tool for locating expertise for software engineering collaboration, which was called Expertise Browser. Because of obvious limitations in these systems (e.g., focusing on certain document types), academia and industry have paid much more attention to systems that are able to mine heterogeneous sources of expertise evidence accessible in an organization. The P@noptic system developed by Craswell et al. (2001) is one of the pioneers of such systems, which proved usefulness of expertise location based on heterogeneous sources. As a result, at the Text Retrieval Conference (TREC) from 2005 to 2008, an expert finding task was launched as a portion of the Enterprise Track. It provided expert finding systems researchers with a public platform to evaluate approaches and algorithms developed for expert finding. Consequently, expert finding received significant attention from the IR research community, and quick progress was made in modeling, algorithms design, and evaluation. (Husain et al., 2019)

Expert finding systems were previously used in a broad range of both domains and environments such as enterprise (Balog, Azzopardi and De Rijke, 2009 and Craswell et al., 2001), question answering communities (Neshati, Fallahnejad and Beigy, 2017 and Liu et al., 2013), an online knowledge community (Wang et al., 2013 and Huang et al., 2017), social networks, and academia. Experts location systems have several benefits (Ehrlich, 2003):

1. Individual benefits — such systems, besides completing their direct purpose of finding experts, also connect people who are unable to interact in person otherwise. They also might improve the individual’s understanding of the processes within the organisation and enable people to customize their responses depending on people they may meet in future. (Husain et al., 2019)
2. Organisation level benefits — expert finding systems help organizations in two ways:

(1) They extract and depict tacit knowledge within organization and help people to have a clear understanding about who knows what; though, data of the expertise location alone is not enough. The organisation benefits in case if this knowledge can be actually applied to support problem-solving.

(2) Any expert finding system must also strengthen the ties between people who are acquainted with each other and increase conversations between people who are not acquainted, thus expanding their social network.

The expert finding task usually has procedures that can be classified into the issues:

1. Expertise evidence (sources of data) selection, which extracts the primary data from which expertise can be located and experts can be identified. Data related to the person should be mined and used in order to confirm whether the person possesses expertise in a given domain.
2. Expert representation — a primary goal of expert finding systems is to supply users with full information enabling them to not only locate expertise but also to select most suitable expert from the sample of the most relevant ones. It is crucial to set a list of criteria to form this list so that it fits system user current needs the most. Generally, an individual might be defined as expert based on his or her documented evidence and social interaction factors (which are also referred to as contextual factors).
3. Model building, which consists of three main stages, which are namely pre-processing, indexing, and modelling.
   1. Pre-processing and indexing is a regular task for document retrieval systems, and a broad range of software tools are developed to solve this problem; for expertise retrieval, though, there is additional challenge. The identifiers of a potential expert (such as names, emails, etc.) need to be mined for within documents, and heterogeneous sources should be brought together to identify experts. Several text pre-processing techniques for traditional IR applications are used in expertise location. Thus, some expertise location systems eliminate stopwords from raw document data in order to increase precision of the results obtained (Balog et al., 2012). Heterogeneity of data sources used leads to specific challenges for both information integration and indexing. For instance, different names may refer to the same person or the same names in different sources might refer to different persons. (Husain et al., 2019) Named entity recognition (NER) and disambiguation play a fundamental role in most operational expert search systems (Balog et al., 2012). One of the ways of conducting NER is to use the expert identifiers as the query and associate the retrieved documents with that person (Petkova and Croft, 2007).
   2. Modelling and retrieval — modelling and retrieval are used for developing models to link appropriate experts with a user request and to rank results based on particular criteria (Balog et al., 2012). In the literature, there is a range of methods for building expert finding models, such as probabilistic, network-based (also referred to as graph-based), voting, and other models (Balog et al., 2012, Lin et al, 2017).
   3. Models evaluation — the expert finding system is an IR system which is put to work with a request or query from the user to the system, to which system responds with a ranked list of experts that should be relevant to the initial query. Usually, expert finding system efficiency is evaluated using a test collection (datasets). (Husain et al., 2019)
4. Interaction design — the visualization of results obtained through using expert finding model is an issue of high importance. Just a number of names sometimes is not enough to as the user does not have the information to determine the relevance of an expert for solving particular problem. That is why expert search result pages usually include a list of documents, conferences, and journals, which help user to judge upon expert relevance for his request. Moreover, including contact details for each expert in the resulting list enables communication (Balog et al., 2012). There is a number of studies Also, photos of experts are essential because users may want to understand the possible seniority or familiarity of an expert before contacting him or her. Thus, the user may be searching for an expert of the same age or level. Moreover, information that might be used to prove expertise is also of significant importance to user when selecting the most suitable person from the list (Balog et al., 2012).

Usually, there are two major expert finding systems objectives — (1) finding a particular expert in a given domain and (2) expert profiling describing the area of expertise of each expert. Earlier studies (Balog et al., 2012, Lin et al, 2017) confirmed that these objectives are two sides of the same coin, since, in combination, they lead to a ranking problem, wherein the collection of the objects that had to be ranked were prioritized. (Husain et al., 2019)

The expert finding systems have been applied to several domains such as academia, enterprise, and social networks. For example, Husain et al. in the systematic review of expert finding systems (2019) found that such systems have been used in academia to find supervisors, paper reviewers, and research collaborators. In this SLR, the authors noted that 68 studies developed expert finding systems. After analyzing all these studies, the researchers observed that these systems were developed in six different domains for different purposes. These domains are enterprises, academia, knowledge sharing communities (KSCs), medicine, social networks, and online forums. Table describes these domains and studies in further details. From this table it can be seen that academia is the largest domain, comprising 44 studies (65%). A majority of the expert finding systems were used in the academic domain, moreover, in this field, the majority of such systems were created for particular purposes such as industry and university collaborations or finding research collaborator, paper reviewer, similar expert or thesis supervisor. Enterprise is the second largest domain including 10 studies (15%). A majority of the studies in this field developed and used expert finding systems for general purposes, except one reference (Paul, 2015), which applied the expert finding system for offering formal help. In another study by Wu et al. (2011), an expert finding system was used for developer recommendations. Other domains are minor and do not have significant importance for this research so we will not explore them in detail, though we still include them in the table for general understanding of this research line.

Table 3. Expert finding systems research papers distribution by domain and task

|  |  |  |
| --- | --- | --- |
| Domain | Task | Frequency |
| Academia University | Research collaboration | 1 |
| Paper-reviewing | 2 |
| University-industry collaboration | 1 |
| Supervisor finding | 1 |
| Similar researcher | 2 |
| No specific task | 6 |
| Academia social network | Research collaboration | 12 |
| Paper reviewing | 7 |
| Finding group of experts | 3 |
| R&D | 2 |
| No specific task | 8 |
| Enterprise |  | 10 |
| Medicine |  | 2 |
| Knowledge sharing communities |  | 6 |
| Online forums |  | 2 |
| Social network |  | 4 |

Most important expertise sources are publications (which are used in 22 research papers), social networks (10) and homepages (8). Some seemingly important sources, especially for some particular reasons of expert finding (e.g. supervised theses for searching for supervisor finding) are clearly underused in existing expert finding models.

Currently used models in expertise retrieval systems are based on three main components of analysis, which are experts, documents and the domain and can be classified into several major groups.

**Generative probabilistic models**, which evaluate the relationships between domains included in the query on one hand and individuals on the other as the probability of the particular expert to have sufficient expertise in the given area. This type of model is the first edition of the Text Retrieval Conference (TREC) Enterprise track, and is considered a popular approach for expertise retrieval. (Husain et al., 2019) The ranking process of experts is based on the probability of person e being an expert on query q (P(e|q)).

Such models apply specific methods, for instance, document-based methods (also referred to as query-dependent methods) which first find the documents relevant to the query domains and then rank associated experts based on both document degree of association with domain and the degree of an expert association with the document. Therefore, experts are ranked based on a weighted set of documents. Examples of such methods, applied for estimating similarity of two researchers in (Gollapalli, Mitra, Giles, 2012) are:

* Okapi BM25 (OKAPI), which uses a vector of terms associated with researcher’s content to represent the researcher profile and considers one such profile as the query and the other as the document to measure similarity between them;
* KL Divergence (KLD), using a probability distribution to represent a researcher profile;
* Probabilistic Modelling (PM), using Latent Dirichlet Allocation as a tool for topic modelling in unsupervised clustering of data and exploratory analysis;
* Trace-Based Similarity (REL), using density matrices to represent vector subspaces for modelling concepts.

**Discriminative probabilistic models** are based on conditional probability distribution instead of joint probability distribution used in generative models. Generative models are strongly based on more assumptions, while discriminative “prefer to let the data speak for itself” and tend to have a lower asymptotic error. (Husain et al., 2019)

**Voting models** use different sources of information separately to rank experts and then combine these ratings by considering them as “voting” for experts instead of aggregating the values directly.

**Network-based models** use graph-based methods based on expert networks. There are many algorithms that are commonly used in network-based models, such as HITS and PageRank algorithms and random walk propagation algorithms. (Lin et al., 2017) Experts and documents are viewed as webpages, and hyperlinks can be seen as the expert–expert or expert–document relationships of the expert finding model.

# CHAPTER 2. RESEARCH GAP AND METHODOLOGY

## Research gap

Expertise location is the complex process that can be used for a wide range of purposes and in order to be useful it has to be highly customizable while being simple to use. Even more than that, each request for help from an expert is unique, which means it is absolutely necessary to provide user with all required tools to specify his query as detailed as possible.

At the same time currently developed methods are created for solving one particular problem: searching for academic supervisor or similar researcher or paper reviewer. At one hand it indeed allows user for easier setup of the system, since all of the users have more or less the same purpose of using it which allows to simplify both user experience and system itself. It is also especially helpful if we consider a case, where such system user does not have precise understanding of which exact expertise he is searching for, since he has to specify only a few basic conditions of the search.

On the other hand, even if we consider a single purpose method, for instance, looking for co-researcher, we would notice that in different cases people are looking for different skills. Sometimes they would like to collaborate with a person that has similar research interests, sometimes they are looking for help with a specific methodology, no matter of the area of interests and sometimes they have personal requirements towards the collaborator they would like to specify, such as professional experience.

If we will further consider that there is not a set of distinctive goals, but rather a specter of unique queries to find an expert, it becomes evident that user should have the ability to select which areas of expertise are important to him and what exact proof of expertise he is mostly interested in.

Existing research line does not provide such multi-purpose method and moreover all methods have only a few source of expertise proof (such as publication or citation data), mostly one or two. This obviously does not give much flexibility and even if a user would want to use that method for another purpose, he might be misled since for this particular purpose those expertise sources that are built-in the method might turn out to be completely irrelevant.

This makes it a widely spread practice to contact colleagues and use personal network in order to find an expert which is less effective, more biased and resource consuming way to solve this issue.

Therefore, creation of multi-purpose, flexible and highly customizable expertise location method is clearly eliminating a research gap.

## Research goals and objectives

The goal of this master thesis is to develop an expertise location method for universities allowing to solve various user problems through customization and using multiple relevant sources.

To achieve this goal, we have to answer following research questions:

* Which main tasks in the universities require expertise location?
* What are currently used methods and techniques of expertise location?
* Which techniques are applicable to universities?
* What method can be used for expertise location in universities?

Thus, we can formulate several objectives that we have to achieve:

* To identify main processes in the universities that can be enhanced by introducing expertise location method (e.g. searching for co-researcher or academic supervisor, educational programs development, etc.);
* To list and analyze known methods and techniques of expertise location process and estimate their applicability to the processes in the previous step;
* To develop a method for expertise location, applicable for any academic environment, considering existing methods and techniques;
* To demonstrate this method on experts from the same university and verify adequacy and precision of results

## Design science research methodology

This thesis uses qualitative research methods by adopting design science research methodology developed by Peffers et al. (2007), which implies six consecutive stages and possible iteration if necessary (figure 1) in order to create an artifact to solve particular problem. In this thesis we aim to develop a method to enable effective expertise location within organization, which explains the choice of methodology.

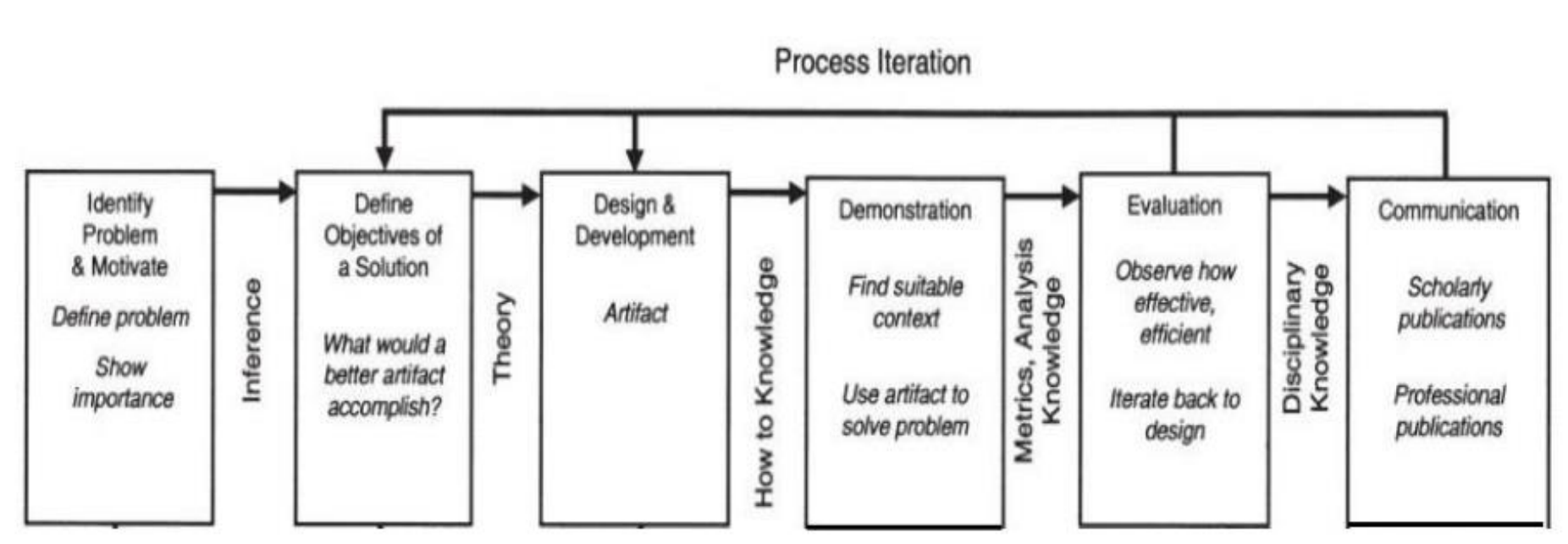


Figure 1. Design science research methodology (Peffers et al., 2007)

Detailed description of how this methodology is applied to this thesis is summarized in the following table (table 4):

Table 4. Methodology of the thesis

|  |  |  |
| --- | --- | --- |
| **Step** | **Description** | **Methods** |
| **Identify problem & Motivate** | Define problem and show importance | * Literature review to identify limitations of existing models; * First round of semi-structured interviews to confirm and update problem definition from literature review |
| **Define objectives of the solution** | What would a better artifact accomplish? | * First round of interviews to understand relevant user needs, expectations and restrictions the proposed method should meet |
| **Design & Development** | Creation of artifact | * Listing tools from existing research * Selection of the most appropriate tools based on interview insights * Development of expertise location method based on the list of existing tools and interview insights |
| **Demonstration** | Find suitable context.  Use artifact to solve problem | * Manually gather and analyze expertise sources for several experts * Provide output for several method usage purposes |
| **Evaluation** | Observe how effective, efficient | * Second round of interviews to evaluate expert profiling precision |
| **Communication** | Scholarly publications | * Suggest further development of the method based on previous step * Suggest further development outside the scope of this research: e.g. automation of the process and improvement of visualization of results |

## Data collection methods

Theoretical basis for the framework development is derived from literature review and analysis of existing frameworks to create knowledge maps and to locate expertise within organization. Information about selected experts for personal expertise profiling is obtained through both publicly available Internet resources, such as aggregators of publication data and personal webpages on the one side and personally disclosed data, for example, about professional experience.

Another crucial source of information, crucial for method development, is conducting semi-structured interviews with representatives of different groups of potential users: students, university management, researchers and experts, which are subjects for the expertise location and expert profiling processes, for confirming or disapproving preliminary needs, expectations and requirements towards method and its output deriving from literature review.

# CHAPTER 3. METHOD DESIGN AND DEVELOPMENT

## Purposes of expertise location and valid sources of expertise

In this and the next chapter we extensively use the results of the first round of semi-structured interviews, during which 11 respondents were interviewed, 7 of which are GSOM employees or students and 4 from other universities to confirm the applicability of the method to other universities. We also combine insights from these interviews with the conclusion derived from literature analysis.

Table 5 contains information on different respondent groups and their preliminary needs and requirements towards method, and Appendix 1 has all the main questions asked during the interview. Also it should be noted that both university management and active researchers are also experts to be evaluated using this method, so they were asked these questions as well.

Table 5. Main groups of respondents and their needs towards the method

|  |  |
| --- | --- |
| **Group** | **Preliminary needs for locating expertise** |
| **Expertise location output user groups** | |
| **Researchers** | **Finding collaborator** for conducting particular research  **Finding expert for paper reviewing**  **Finding speaker for conference** |
| **Academic directors** | **Understanding overall organization expertise** for understanding capabilities for creating new and developing existing courses  **Assigning professor** that fits particular course best  **Understanding gaps** in overall organization knowledge |
| **Students** | **Finding academic supervisor** for writing thesis in case of the significant difference of desired topic from those offered by potential supervisors during their presentations |
| **Respondents, which are subject to expertise location process (are also users at the same time, so are interviewed from both perspectives)** | |
| **Experts** | **Precise expertise estimation** leading to proper decision making of other groups of users based on the method output |
| **All the groups, since anyone might have a need to find expert to consult with over specific question or issue** | |
| **Common for all groups** | **Consulting with best available expert on particular question or issue** |

Distribution between groups was the following: 2 respondents represented university management, 2 were students, 2 were experts, not active in research area and the last 5 were active researchers, which are presumably the most relevant users of the method created using the method we seek to develop.

As for interview results, university management confirmed the relevance of both need to find the most suitable expert, for instance, for participation in public activities to improve university visibility and brand or for lecturing and need to have an overall organization expertise overview.

Students have confirmed the lack of tools to select the most appropriate scientific advisor for topics not listed in potential supervisors’ presentations. At the same time currently used system of allocating students for writing theses can hardly be fully integrated with suggested method.

Experts that are not active in conducting research, are mostly looking for experts not in academic environment but rather within professional communities, but they have stated their willingness to use the suggested method in case if it is simple to use and provides useful results.

Active researchers confirmed the needs to find experts for collaboration and paper reviewing, but also mentioned such purposes as finding speaker for the conference organized.

As for the sources relevant for each of the goals, respondents agreed that expertise classification into three major areas is relevant and confirmed the main sources that are used in this research: publications, conferences. citation data, positions held, courses taught. Some of the sources were mentioned by some of the respondents but doubted by others, such as grants and age.

## Requirements for the method of expertise location

Interview also revealed several important requirements proposed method should meet in order to be considered useful by the potential users, which are listed in Table 6. There are also some requirements that can be drawn from literature review and interviews irrespective to the group.

In most cases, university staff currently prefers personal connections to find experts due to the fact that they are acquainted either to an expert or a person who can suggest an expert in almost any field. That means that in order to be used in practice our method has to be simple to use and have to be trusted in the degree comparable to recommendations from colleagues.

There is already a number of documents that require candidate experts have to fill in to provide university management with relevant information about their scientific and professional activities for their personal profiles or other purposes, so they would like to avoid reentering the same data several times if possible.

Since workload of experts varies greatly as well as workload of one expert over time, interviewed experts stated that they would like to have an opportunity to select their status, indicating their willingness to collaborate at the moment.

Table 6. Main groups of respondents and their requirements towards the method

|  |  |
| --- | --- |
| **Group** | **Preliminary requirements method should meet** |
| **Expertise location output user groups** | |
| **Researchers** | **High level of detalization** to choose collaborator whose expertise covers most relevant domains required for this research that searching researcher lacks  **Personal fit with researcher looking for collaborator:** age, overlap in expertise, complementary expertise  **Preferred methodology match:** qualitative or quantitative, etc. Since some researches might be based on the methodology original author has no much experience with, he might look for expertise in this area, no matter what the collaborator research interests are |
| **Academic directors** | **Possibility to compare experts with external ones** to gain insights about competitive advantages of the university  **Possibility to locate missing expertise,** not only within domains there are internal experts  **Multiple domain expertise location**, since one course might be cross-domain  **Data relevance and precision** |
| **Students** | **Common research interests.**  **Preferred methodology match:** qualitative or quantitative, etc.  **System could advise** most suitable **domains**, based on student personal interests, as well as available supervisors’ ones.  **Integration with the official system** of assigning supervisors |
| **Respondents, which are subject to expertise location process (are also users at the same time, so are interviewed from both perspectives)** | |
| **Experts** | **Minimal required time and self-disclosed data** to keep expertise location results relevant and precise enough  **Data security**  **Minimal personal involvement**  **Presence of the tools to interact with own profile** |
| **All the groups, since anyone might have a need to find expert to consult with over specific question or issue** | |
| **Common for all groups** | **Expert current availability**  **Automatic domain identification from the request**  **Distribution of requests** to not overload most knowledgeable experts  As solution sometimes lies in the other domain, expertise in related fields should be taken into account |

Since the method is designed for academic environment, it should take into account peculiarities of such environment.

As far as users of the method are supposed to be a part of academic community they often are interested not only in the results of such a method, but also in proof of its reliability, which means that method should possess a certain transparence.

There is a broad range of possible purposes to find expert, as well as a large number of domains of knowledge, therefore method has to provide powerful customization options.

Most of the expertise sources are stored at the same online resources that should be outlined and parsed for information: e. g., citation and publication data and personal homepages at faculty website.

## General concepts of the method of expertise location

The general idea of the method is to analyze input data of different types and to transform it into specific outputs as shown in the figure 2.

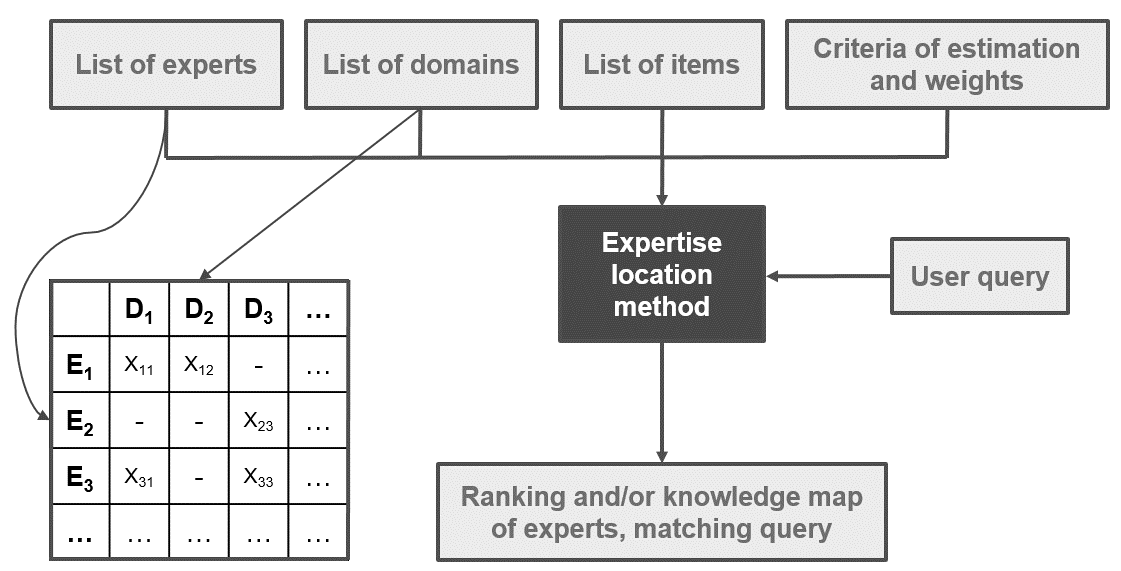


Figure 2. Inputs and outputs of the method

In this research in order to develop and apply the expertise location method we use as a general framework the ideas of general probabilistic model “Model 2” proposed by Balog et al. (2006), which implies that degree of association between expert and query can be calculated as weighted sum of each document relevance to a given query, using as weights degree of association between expert and the document. We use this logic to estimate minimal grain of method results, namely the expertise of a particular person in a given domain obtained from the given expertise source (figure 3).

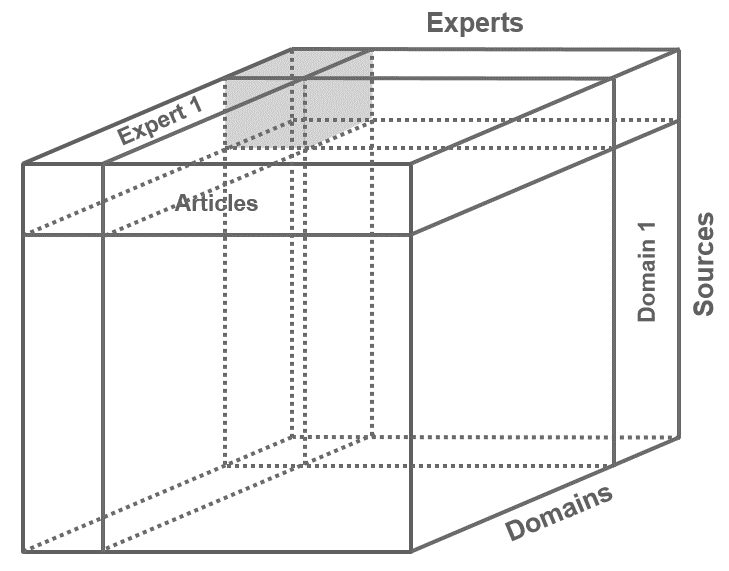


Figure 3. Minimum grains of data preprocessing output

This idea is also used when we aggregate these grains into the final results to be visualized: we weight such grains by their relation to query.

Query defines both weight in terms it has to specify the target domain(s) for which we are estimating expertise of an individual and the goal of the search, such as searching for academic advisor, thus indirectly setting the weight of each source.

We also use journal ratings in Scopus and Russian Index of Scientific Citation (RISC) in order to evaluate how reputable journal is along with checking whether the journal is included into the list of peer-reviewed journals in which the main findings of dissertations for the degree of candidate of sciences or for the degree of Doctor of Sciences should be published, which issued by Higher Attestation Commission.

We will also have some steps in method as “black boxes”, that take a given input and provide specific output. This is done due to several reasons:

* currently used tools are interchangeable and constantly developed, which means that we have to ensure that we provide the input in the right form for these tool to process and the next step to be able to take the output of the preceding black box as its own input, so that any algorithm could be used within this stage;
* list of sources should be defined after detailed analysis of the given organization current state and needs, that can be addressed by implementing expertise location method and therefore shape the whole method, so each step should be completed using the most appropriate techniques;
* most textual sources retrieval techniques have to analyze the full text of the publication which might be unavailable for free, which will put additional restrictions on this method implementation.

## Method for expertise location

The method, as it was already stated, consists of consecutively applied tools and techniques that allow to transform raw data, partially available online and partially uploaded by candidate expert into expertise indicators and personal profiles of each expert.

Overall idea of the method is that following several steps and based on particular needs of expertise seeker on one hand and available expertise sources on the other tacit and fuzzy mutual understanding of colleagues’ expertise would be replaced or at least supplemented with formal map of knowledge within organization and ratings of experts for each particular domain.

The method includes following steps:

1. **Input stage**
   1. ***Automatic input of the data, that is available online***

At this stage we list sources that are used to obtain following data: publications, citations, courses taught, participation in conferences, grants, applied experience (e.g. consulting projects or position held at commercial organizations) membership in organizations, such as conference committees, editorial boards, etc. After that we retrieve all required data about candidate experts and transform it into a database, having following fields: expert identification number, expert name, publication title, year of publication and, finally, list of keywords, which will further allow us to determine which domains this work is strongly related to.

* 1. ***Manual input of missing data by experts***

This step is conducted by asking the expert to fill two gaps in automatically collected data: firstly, provide data on sources of expertise that are not available online (for instance, if expert does not have personal page or haven’t stated there his memberships), and secondly, ensure relevance of the data from the first step in case of its inferiority by adding missing items.

1. **Analysis stage**
   1. ***Evaluation of each item importance***

Not all items are equally significant in estimating overall expertise of a given expert in particular domain. There are several general measures, which are common for several or even for all expertise sources: language used, relevance based on date, number of collaborators. Not all of them are used in the demonstration case for simplicity sake, however. Moreover, each expertise source has its own measures of relevance for expertise retrieval process. For example, for publications, it is the rating of the journal, for courses taught it is the level of education, lectures are aimed for: bachelors, masters, etc.

We estimate these measures in the range from 0 to 1 and then multiply them to get item weight as a single measure:

,

where are different measures of the same item.

* 1. ***Listing related domains and measuring this relation***

From the verbal description of the item (title and keywords of publication, title of position and description of responsibilities) we select the most related to this particular item domain(s) from the domain structure and assess relationship between them as 1 in case of clear match, 0.5 in case of high likelihood and 0 when relation does not exist. Structure of domains is defined manually or obtained from public ontologies. In the demonstration case we used structure developed by internal experts

* 1. ***Integration within each expertise source and each domain of knowledge***

After we have estimated each item weight and its relation to different domains, we are able to calculate each domain overall expertise. Proven by a single source of the particular expert by calculating weighted sum of all items from a single source, using the following formula:

P(E|d) = ,

where shows weight of the given item and shows relation of the item to particular domain.

The output of this step is a matrix, having three dimensions: expertise source, expert and domain, having aforementioned sums, normalized by dividing by the maximum value of any expert expertise in any domain extracted from the same expertise source, as values:

* 1. ***Aggregation sources towards expertise areas***

Summarizing different sources is conducted within two steps. Firstly, we estimate expertise in three major areas of activities: research, teaching and applied experience. We suggest that expertise in research should be based on publications and citation data and, to less extent, conferences, grants and membership in academic organizations, such as editorial boards and organizational committees of conferences. Teaching expertise is calculated through courses taught, courses developed and theses supervised, while applied experience is estimated through positions held at commercial organizations, consulting projects and other experience, that can characterize practical application of expertise.

We apply the following formula for this step:

),

where SW are source weights, and EA is an estimation for expertise referring to a specific expertise area (research, teaching or applied experience) of a given expert in the particular domain

At this stage we have data with three dimensions as the output: namely, experts, domains, and areas of expertise.

1. **Visualization stage**
   1. ***Aggregation of the data according to query***

After that, we finally aggregate data according to the query into two-dimensional structure. This can be done in one of three ways, according to query:

* **experts — domains.** Supposedly the most common way to present results, since it allows to rank experts by their overall expertise in a given domain. We eliminate areas of expertise dimension in this case by calculating weighted sum of their values, based either on preset of weights, associated with the goal of the expertise search, which is specified in the query, or weights, directly defined by the user:

,

where are query weights, are values of the expertise of the given expert in particular expertise area in the given domain that we sum for each expert and domain and n is number of the expertise areas (in our case n = 3).

* **domains — areas of expertise.** This table aims to show an overall picture of expertise within the organization and is supposed to be useful mostly for university management by helping to locate gaps in knowledge and particularly strong areas, which can be used as competitive advantage through introducing new courses and developing current programs.

This way of dimension reduction is carried out through calculating simple average values:

, where n is the number of experts.

* **experts — areas of expertise.** This format of data allows to get an overview, which experts are more research-oriented, and which put more effort at teaching, to potentially introduce personal motivation system:
* , where n is the number of domains.
  1. ***Building visual knowledge map, depicting the results*** (when applicable)

The results are to be visualized using knowledge mapping technique, which makes output easy to grasp and use them effectively.

# CHAPTER 4. METHOD DEMONSTRATION

## Description of organization and environment

To demonstrate the practical application of the developed method the case of Graduate School of Management (GSOM) was used. GSOM is a leading business school in Russia, which was founded in 2005 by merging Faculty of Management, the Research Institute of Management and the Special Faculty for Personnel Retraining. In just 15 years it has 7500 graduates, 1200 students, 200 corporate partners which are involved into the development of curricula, give lectures, invite students for internships and organize joint projects, 78 international academic partners, 6 international research centers and 3 international organizations.

It also is regularly included into the world rankings of educational management programs and international rankings of business schools of American and European rating agencies.

Since 2013 GSOM Master in Management program has been continuously included in the Financial Times Global Masters in Management ranking and by 2017 improved its position by #23 out of 90 best Master in Management international programs. Rankings of the international business newspaper Financial Times, published in more than 20 countries around the world, are a generally accepted indicator of the quality of a business school or a program.

All this information has direct impact on creating expertise location method.

Firstly, GSOM has tight international relationships and provides some educational programs in English and at the same time the majority of employees and students are Russian, which implies, that expertise sources both in English and Russian, and for some purposes of expertise location the level of English of an expert plays an important role.

Secondly, GSOM has close relations with corporate partners, which might be interested in using the method to find the most suitable expert. That means that such corporate partner employees as a system user might have non-academic background and therefore lack deep understanding of knowledge domains structure and expertise proof. This calls for a simplicity in the method usage, allowing anyone to use it effectively.

Finally, existence of the research centers, partnerships with corporations and claimed focus on preparing the national elite of managerial personnel proves that all three major areas: namely research, teaching and practical application of expertise are crucial for both management and community of the school.

To conclude with, GSOM website has personal pages of each lecturer, which can be used as primary source of the data for the method, which can be complemented by independent Web resources, that contain publicly available publication and citation data.

## Identifying relevant goals of expertise search

The primary purpose of expertise location method is to provide users with a relevant and easily understandable response for a specific request: which can be, for example, ranked list of suitable experts or knowledge map.

That means that in order to build effective method we have to develop a clear vision what are major goals of expertise search that are relevant for universities and particularly GSOM. We have conducted several semi-structured interview with representatives of different groups of method users.

The first group consists of researchers, that have claimed searching for collaborator or paper reviewer as major purposes of expertise location. Currently this is usually done through personal contacts and recommendations of colleagues, which is effective due to the relatively small number of researchers in a single academic research domain, as it was mentioned by one of respondents. Therefore, developed method should be effective and simple enough in order to be used to complement currently used approach. On the other hand, in the case of cross-domain research, finding a suitable expert through personal relationships is significantly more challenging due to the fact that the links even between different faculties of the same university are weak, experts in different domains attend different conferences and events and therefor have little chances to get acquainted to each other, which makes the system potentially useful for locating expertise in “distant” knowledge domains.

The second group of method users is university or faculty management. They would like to have, firstly, an opportunity to find the most suitable expert either for participating in public event or leaving a comment on particular issue in the media in order to increase university and/or faculty visibility or for developing the new course or giving lectures and, secondly, the overview of overall expertise of the organization. Such tool will enable them to:

* identify knowledge gaps to eliminate them and locate the strongest domains, which can be perceived and used as a competitive advantage;
* evaluate capabilities in developing existing courses and programs and creating new ones;
* be aware of risks connected with the concentration of expertise in a single domain in a few individuals. In this case their leave would significantly damage overall organization expertise and might potentially undermine the quality of several courses or even programs, whereas having several experts in each domain helps mitigate this risk

The last major group consists from students, who are mostly looking for an expert when choosing research advisor. In GSOM, though, there is fully designed process of allocating students to potential advisors, which includes presentations from professors, where they provide information on their research interests and preliminary topics of theses. But even in this system there might be a case, when student would like to suggest his own research area and does not know whom to approach for the discussion about possibility to collaborate on such thesis. Moreover, current advisor might lack expertise in some domains thesis is related to, so expertise location method could be used to suggest a suitable expert to consult with.

## Gathering primary data

Primary data consists of several major parts: list of experts, domain structure, criteria for estimation and weights and, finally, list of items. User query can also be viewed as the input of the method, but, it is unnecessary to run expert profiling process irrespective of particular expert searching purpose and without queries method still generates useful data, that can be used for default requests. This is why we can define it at the stage where it is required for customizing final output of the method.

List of experts to estimate is provided by the university, including lecturers, scientists, not involved in teaching process and even doctoral program students, depending on university policy, budget, time and human resources available for the project and goals it seeks to achieve by implementing the method. For the demonstration we have profiled three experts within GSOM.

Domain list can be either developed internally by experts based on organization management requirements to the output or integrated from the outside, in case, if any of the existing hierarchies or classifications are suitable enough. For this thesis, there already was internally developed domain structure of GSOM expertise domains, which was used “as is”. It consisted of 12 domains, each having 11-20 subdomains. Some of subdomains belong to several domains simultaneously (Appendix 2).

Criteria for item estimation and weights are to be primarily defined by the organization policy and/or interviews and surveys conducted among system users and experts. They can be updated regularly either automatically or manually based on feedback or changes in policy.

Items, which are the primary subjects of analysis, are gathered from several sources. Most of it is stored within personal homepages of candidate experts at GSOM website: list of most important publications, conferences, positions held in corporate sector, conferences participation and courses taught.

Then the information collected is enriched by using additional sources, such as Pure, which is a service that contains publications and citation data on all the lecturers of Saint-Petersburg State University, and particularly GSOM, as well as his or her h-index, ongoing and finished projects. It also provides the results of analysis in the form of “fingerprint”, which is a set of most relevant keywords, characterizing research direction, but these results are far from accurate which might be caused by wrong association of works with domains, inaccurate domain structure and lack of effective weighting sources by their importance for expertise profiling. Another website that contains data on publication list with citation data is eLibrary, which is Russian digital library, integrated with Russian Index of Scientific Citation. Moreover, this resource has a wide range of metrics that could be potentially integrated into the method, such as average impact-factor of journals the author was published in and impact-factor of journals citing author.

After this step information should be confirmed and updated by the expert manually, since, for instance, consulting projects and research in progress is not available online and some items from sources that are generally accessible online as stated above might be missing and therefore should be added.

Another important piece of information that we used within the method is journal rating, based on Russian Index of Scientific Citations (RISC) and Scopus ranking, as well as Higher Examination Committee (HEC) recommended journals list.

This list of primary data is not fixed, it can be updated even for the same organization over time and should be completely reassembled for any other university willing to implement the method. Though no changes might be needed at the end, initiating and conducting reevaluation of the sources should be done for this method to be as effective as it could be.

After all these steps the information should include items from following sources:

* **Articles.** Each item should have title, date of publication, number of coauthors, volume and keywords. Based on this information we can estimate date-based publication relevance, journal rating, personal impact, and degree of relation to different domains. All of these coefficients should be in range from 0 to 1.

Using the following formula:

where is impact of the item towards total expertise of a given individual in a given domain,

*TR* is time relevance (in our method we considered works that were published less than two years ago are assigned TR equal to 1, less than 5 years equal to 0.8, less than 10 years equal to 0.5, less than 20 years to 0.2 and older articles receive TR equal to 0.1).

*JR* is the journal rating, which is calculated in a more complicated way, based on inclusion in mentioned above HEC list and rating in Scopus and RISC using the table X,

*PI* is personal input based on number of authors (this coefficient for works, written by a single author is equal to 1, with one coauthor — 0.7, works having 3 or 4 authors — 0.5 and for more authors the coefficient is 0.3).

*DR* is relation to domain. Although the most popular technique to estimate that is to find most commonly used terms within the full text of the work and to compare them with the full list of domain-specific terms, this approach requires the list of such terms for each domain, created either manually by the experts before implementing this method or automatically derived from the domain-specific literature, proposed by the experts. For that reason, combined with the need to have full texts of all analyzed articles which is an additional restriction of such technique, we simplified the calculation of this indicators to be based only on title of the work and keywords. If the title of the work allows us to assign it to particular domain(s), we take DR equal to 1, if from title we cannot do this, but keywords help us to do that, we take DR equal to 0.5.

*VW* is based on the volume of the work measured in the “printed pages”: works that account for more than 1 printed page are granted 1 for this criterion, from 0.5 printed pages – 0.7 points and lesser works get 0.5.

Such simplified approach gives us the result from 0 (unrelated to domain) to 1 (article published less than two years ago in the top-10% Scopus rating journal publication with volume more than 1 printed page, written by one author and title related to a given domain). Lowest non-zero value so far would be 0.000075, which in practice allows to take into account only valid proof of expertise;

* **Studies.** Each study has to have the same attributes as the article, except for journal rating, which is inapplicable in this case;
* **Conference papers.** Same attributes as for studies;
* **Teaching aids.** Same attributes as for studies;
* **Citation data.** Each of the items from the aforementioned sources has also the attribute of number of citations, which multiplied by summarizing indicator gives us the understanding of recognized expertise of the particular expert in the particular domain. It can be further improved by taking into account the level of citing publications;
* **Positions held.** Important factors that help us to estimate expertise through the lens of this source are relevance based on date of retirement from position and total duration of holding this position. We estimate coefficients that are based on this data, as follows:

For duration of time in the position we set 1 as the value for more than 10 years of experience, 0.9 for more than 5 years, 0.7 for more than 2 years, 0.5 for more than a year, and 0.2 for less.

For the relevance we use the same classification as for the articles and other sources that have the same attribute.

Obviously, degree of association with a given domain should also be included into the formula as the last factor. Basically we give 1 if from the position title we can derive clear relation to domain and 0, if we cannot. However, his source should be carefully examined by the expert and each position should be connected with all the related domains and the degree can be any value from 0 to 1;

* **Consulting projects** should have relevance and degree of association with given domain(s), calculated the same way as for position held and one special binary variable taking values 0.5 and 1 depending on whether the project was pro bono or paid.

All these items should be listed within a single database, having all these indicators, name of an expert and domain as fields. Each item can be listed several times for each domain it is associated to and for each author within the organization. For instance, the article, which can be associated with two domains has four authors, three of which are GSOM employees and therefore candidate experts for the method, should be listed 6 times (twice for each expert). All the other fields do not change in this case. Moreover, additional 6 rows are added for such article as it is viewed also as citation source.

At the end of the data collection and preprocessing we should have a database that can be analyzed using SQL requests, Excel formulas or BI advanced tools depending on the organization resources and policy.

## Data processing and intermediate results

Having a single database with three dimensions, which are namely experts, domains and sources filled with values from 0 to 1 showing the expertise of a given expert in a particular domain proof demonstrated by specific expertise source, we are able to aggregate it by calculating a simple sum of all the values within each “cell” of this structure (related to the same expert, domain and expertise source).

However, this would lead us to values which can be any number above zero. If an expert has one hundred publications with average summarizing indicator of association to a given domain equal to 0,5, his score would be 50. It is impossible to make any judgments on whether it is much or little until we compare it with other experts. That is why after calculating sum of all items in the same “cell” we have to normalize them to be within range 0 and 1, which will enable us to understand whether the person is expert in this domain or not without comparing it with other values and therefore will make the data easier to grasp.

Another crucial point for normalizing values is that for different sources the maximal values would be very different. Individual can possibly have even several hundreds of publications, but he definitely cannot have comparable number or positions, so without being normalized to be measured on the same scale different sources could not be aggregated into a single indicator, which is necessary step within this method.

To normalize values, we will divide them by the largest score for the same source between all experts and domains. This will allow us to compare different experts and different domains, as before this transformation of values, but make further analysis possible and results easier to understand. The only drawback of this method is that it doesn’t allow for external comparison. For instance, if almost no one in the organization attends conferences, then it might be possible we take a very little value as 1. From one perspective, it is useful, since it makes the differences between expert more visible by setting the scale based on maximal differences within organization, but, on the other hand, it will not show organization its weak places if we will study two-dimensional matrix of experts and sources. In other words, it is impossible to state whether the experts lack professional experience since at least one of them will have value of 1 in at least one domain.

However, this issue is possible to solve only if we know, which value is theoretically largest possible. We can assume that only for several sources, such as position held (we can take minimal legal age to be employed and maximum life span and use this to calculate highest possible pre-normalized value, by considering that person had only positions directly related to a given domain and changed them after required number of years to maximize cumulative value), but this is useless, since unachievable in reality. Other sources are impossible to estimate, so the question is always to find the most highly ranked expert. This brings us to the fact, that having more external experts listed in such method will increase the probability that such highest rated expert will be within the list.

However, this aggregation and normalization is not the end of analysis. The next step for estimating overall expertise of individual in a domain seems natural: reducing number of dimensions to two by weighting sources. Indeed, that will give us the single expertise indicator, which we can use to rank experts and to estimate organization knowledge gaps in particular domains, but setting all sources equal might be misleading.

For instance, if an expert is outstanding lecturer that developed several courses and teaching aids but has poor results in research activity he would have an average score and be ranked higher than mediocre researcher in the same domain, which has almost no teaching activities. For users of the method that are looking for co-researcher this ranking would clearly be useless and, when trusted, even harmful.

This calls for that weights for aggregation should be somehow defined by the method user based on the purpose of finding an expert. There are two extremes for solving this problem:

* precise direct tuning, which might be too complicated, especially for users that have non-academic background and for the most effective usage of the method it is crucial to clearly understand the way it is working, which calls for transparency of the method and required time to examine it;
* simple selection of the goal and using preset weights, possibly not even shown to user (e.g. for co-researcher finding we would assign high weights for publications and citations and low weights for lecturing). This approach makes method easier to use, but in case of uncommon goal of searching for the expert, not included in the list of possible purposes, it cannot provide any useful output. Moreover, even within the same goal, there is a range of requirements that can differ from query to query, which is impossible to take into account.

This is why though we use presets for particular goals they can always be easily changed by the user if needed to. But if all the calculations are done only after the query is received by the method, it will take time. In case such system has many users, that delay might be a crucial issue. Moreover, having all settings available, user can get confused and as using system is uncomfortable for him or her, he or she would prefer to use traditional approaches and refuse from using the method.

We have already stated that all the respondents confirmed that there are three major areas of activities relevant for expertise location process: research, teaching and applied experience. That enables us to conduct aggregation of expertise from different sources into a single indicator into two major steps: firstly, we group all sources within these areas using common weights and then user sets proportion of these three areas of expertise required for his specific purpose. Method still should offer some presets for most common goals, but these presets are easier to change and faster to calculate than weighting all the sources for all the experts.

There is also still possibly should be opportunity to manually redefine weights even within these three areas but this tool supposedly would be rarely used and this would still reduce the load on the system and simplify query formulation for vast majority of users.

Predefined weights of separate sources within three areas are selected based on the environment and organization peculiarities and therefore are subjects to organization policy. For the case of GSOM based on interviews we have defined following weights:

* Research activity. 80% of total value is based on publications (30% - articles, 30% - monographs and 20% - conference papers), 10% on citation data and 10% on membership in organizations, such as journal editorial boards and conference organizational committees;
* Teaching activity. 40% of value is based on courses taught and 60% on textbooks and teaching aids written;
* Applied activity. 50% of the value is based on position held, 30% on consulting projects, 10% on membership in organizations and 10% on conference papers.

These could be changed in two ways: manually by experts or automatically based on using feedback on the relevance of output, A/B testing technique and machine learning.

This three-dimensional matrix (experts, domains and areas of activity) is the final output of this stage from which system could extract all necessary information for a given query.

## Query-based final stage of data processing and visualization

As we have already stated there are two key points at this stage: calculating the data to be visualized and visualization of this data to make output as relevant for satisfying user needs as possible. Both of them are defined by the query which is reflection of the purpose of a single user seeking for expert. That is why it is crucial to develop a clear understanding of what information user would like to obtain when using the expertise location system for each of the common goals.

We should state that precise description of user interface and design of output is beyond the scope of this thesis: we aim only to define the list of what should be the output and how it should be visualized, more specific solutions about layout, software, color, etc. are up to the particular organization policy and budget.

* **Finding collaborator or paper-reviewer**. Since these purposes are similar in terms of required competences and desired information. Main things that are relevant for users pursuing these goals are:
  + ranked list of most suitable experts, which is simply top positions of the list of experts in a specified domain(s) that we will get by weighting the areas either as preset suggests, which is 60%–10%–30% (here and further: research activity — teaching activity — applied activity) or custom settings. Note that user should be able to select several domains and set their weights, sum of which should be equal to 1. The table should include 4 columns for each domain (1 per area and 1 aggregated) and in case of the query with several domains specified 1 extra column with an indicator showing the overall degree of matching the query (Appendix 3);
  + opportunity to view expert profiles and most relevant details. Each expert row should be clickable and lead to expert profile, that contains the knowledge map, depicting full ontology of this expert knowledge in all domains, including those that were not a part of a query (Appendix 3). Those which user is interested in are highlighted by different background color in order to show what this expert is best at and how does it match with the query. There might also be a list of the most important items for the query, which calls for consulting with preprocessed primary data before any aggregation, slowing the system. User should also have comparison tool building a single knowledge map for a few experts, highlighted with different colors enhancing decision making process. One of the compared experts might be user himself to better understand expertise in which domains overlaps with his own and in which domains complements it;
  + contact information and willingness to collaborate over particular research at the moment. This information is filled and updated by expert, that might provide phone, e-mail or any other preferred way of communication, areas of interest or, vice versa areas expert is not willing to research at he moment, which might be different for different goals of search and current willingness to collaborate at all. Areas of interest might not correspond to expertise of an individual, so should be input manually.

Experts should also have the option to set different information shown to different groups of users: students, colleagues, corporate partners, external researchers etc.

* **Finding academic supervisor.** Students, using the system are interested in:
  + List of individuals with highest expertise in a specific domain since this purpose is relevant only if student has his own area of interests for writing thesis which does nor overlap with any of areas possible supervisors stated during their presentations.

Suggested weights for this goal are 40%–40%–20%;

* + contact information and how many students can this expert accept to supervise. This information is similar to the last block of information for collaborator or paper reviewer seekers, but willingness to collaborate should be replaced with a precise number of slots for supervising theses.
* **Finding lecturer or developer of the new course.** This is the only common goal that is based mostly on teaching area of expertise using the suggested proportion 15%–70%–15%. Management that is supposed to use the method for this purpose, already has the contacts of the expert, therefore output should include only:
  + Ranked list of the experts matching the query;
  + Knowledge map of the candidate expertise in all the domains with relevant ones highlighted and the most important items. The only difference from the similar profile from the first query type is that for this type of queries items are more obligatory and should be only from teaching area, since this is an important decision to make that calls for a decent proof of one’s ability to develop a course and lecturing experience cannot be successfully replaced by research expertise, especially if lectures are held in English.
* **Looking for overall organization expertise overview.** This is an absolutely different type of query, which should use data without reducing the number of dimensions. That can be explained, that there might be a wide range of insights management might be interested in a part of which might be lost if done so. So system should respond to such query with interactive report in the form of ***competence knowledge map***, highest level of which is the table that should be constructed using experts as columns and major domains as rows and each cell should contain bar chart with a bar for each area of activity and a line of average (Appendix 4).

Each expert should have one additional chart for aggregated expertise, the same should be done for each domain. Top left cell above domains is used for overall organization expertise estimation. Aggregation rules should be customizable by sliders above or below the table, which will move the average lines in each cell.

Each major domain should be clickable showing hidden rows of sub-domains (previously referred to as domains), which on click lead us to the particular sub-domain overview, which is similar to ranked list from the previous outputs but with slider to customize weights of final aggregation stage. Each expert should also be clickable and lead to expert profile also with the ability to customize weights within output. Cells on click should show the list of all items proving expertise of the particular expert in the particular domain, ranked by importance within the table containing primary data.

Such form of organizing output contains all potentially useful information, which is structured and displayed at any level of the map in a user-friendly way, enhancing decision-making process.

The final results that we have obtained for three experts were discussed with them through unstructured interviews and were generally accepted. There were some minor disagreements upon specific areas in some domains, when respondents stated that their expertise was underestimated, which could be explained by the fact that some items were not included in the primary data. This shows that we should aim to include as many items in the primary data as possible. Anyway, expertise estimations were confirmed to be fair overall, and we can improve them even further by rebalancing weights and replacing techniques within stages with more advanced ones.

# DISCUSSIONS

## Benefits and limitations of the method

The developed method has several benefits, which can increase an overall effectiveness of university or even faculty in several ways.

Firstly, it allows for optimization of a significant share of internal processes that require from the particular individual knowledge which he or she lacks but which exist within organization in tacit form. Usage of such method allows for efficient and fast way to retrieve that knowledge by providing seeker with information about who might possess such expertise.

Secondly, experts are required not only for consulting purposes, but also for playing specific roles in particular process: head of program, lecturer for particular course, academic supervisor, paper reviewer. Suggested expertise location method automates the process of creating the list of candidates for such role and enhances decision making process. This leads to both reduction of time resources for assigning the most suitable candidate for the role and improvement of selection process, increasing probability that the most appropriate expert will be assigned.

Finally, such system provides helpful information in the easy to understand way, supporting a wide range of organization management issues:

* human resource management by identifying expertise gaps and risks, related to knowledge concentration within a few key experts and their potential leave. This information might be used for employee training and recruiting purposes;
* development of courses and programs by highlighting domains with the strongest overall organization expertise which can be perceived as core competitive advantage and therefore extensively based upon when developing courses and programs;
* providing consulting or other services for external corporate clients, as well as collaboration with other organizations, which is enhanced through both realistic understanding of core competences organization possesses and selecting most suitable consultant(s) for each particular project, strengthening brand.

Nevertheless, the method has a number of restrictions. The most important one, is that though the method is based on the algorithm, which is based on both theoretical information and interview insights and composed of steps each of which has input and output defined in details, each of the steps can be perceived as black boxes, which can be replaced with any technique of data processing, as long as it takes the same input and provides the same output.

This was done due to several major reasons, discussed in detail in “Frameworks and tools” section, namely constant development of such techniques, specific requirements of the most widely used tools (e.g. access to full text of all analyzed sources) and, most importantly, list of sources and requirements towards method should be defined after detailed analysis of the organization current state and needs, so each step should be completed using the most appropriate technique.

Another limitation of the research is the possible bias towards GSOM case, since the majority of the respondents were representing this university, so the choice of sources, most common needs, preset weights and concepts of the method were derived from the business school environment specificities. On the other hand, this could be overcome during the application of the method to another case by conducting interviews within that organizations to redefine inputs, that are still processable by the method. For instance, if we would apply the method to technical universities, we would be able to use patents as another source of expertise, if we confirm their importance, define how do we estimate each item and how do we include this source to areas of expertise.

For this research primary data collection and preprocessing were carried out manually, since it is more effective for the small sample of experts to be evaluated to test its effectiveness and data collection techniques are outside of the scope of the main goal to create method.

Though method was verified by asking several experts to estimate the precision of their personal profiles and therefore is confirmed to generate fair output, the overview of overall organization expertise generation was conducted in only theoretical way and therefore cannot be validated.

## Suggestions for further research

There is a lack of qualitative studies in expertise location field, exploring effectiveness of application of such systems. Also the method could be adapted for using in any environment by changing primary data sources, setting hierarchy of sources aggregation and shaping output structure. This method itself can be improved through automation of primary data retrieval, usage of the most suitable algorithms within each step and feedback integration to shape preset weights.

# CONCLUSION

The goal of this master thesis was to develop an expertise location method for universities to enhance its processes through boosting personal effectiveness by providing a tool for access to internal tacit knowledge and through supporting decision making process for university management. Through analysis of theoretical background and results of semi-structured interviews the desired outcome was achieved.

The theoretical contribution of the thesis is in developing knowledge management field of study through enhancing knowledge-mapping practices by developing expertise location approach through proposing multi-purpose method allowing to aggregate data from different sources into a single indicator of query-related expertise in range between 0 and 1. Moreover, study examines academic environment for applying such method, being environment-specific, but could be applied to any other domain, if all the inputs and expertise areas are reworked. Method creates a bridge between managerial needs to effectively use internal knowledge assets on the one hand and computer science techniques on the other.

The practical implication is in offering a new tool, relevant for enhancing a major share of all internal university processes and university management, acting like a basis for both expert seeking methods and knowledge maps. It can be tuned to serve any specific need of anyone to find an expert: to consult with, to find collaborator, paper reviewer, academic supervisor, conference speaker or lecturer, to assess overall organization expertise to support strategic decision making by identifying core competitive advantages, critical knowledge gaps and risks, created by high knowledge concentration.

The main limitation of this paper is the validation of the method only for one organization and only for particular candidates’ expertise, which may not be enough to implement it in different universities as is. Moreover, since benefits partly lie in the area of better decision making in courses and programs development it is hard to estimate them after several years, since not only development and implementation themselves take time, but also the results can be estimated only after some delay.

Further research might be aimed at applying this method as framework for different environments, bring more qualitative approach to this field and to test this method by practical implementation it in order to confirm if it is applicable for any university or should be modified. Method could also be enhanced by automatically adjusting presets by introducing machine learning using feedback to automate method iterative development using feedback.

# REFERENCES

1. Afzal, M.T.; Maurer, H.A. Expertise Recommender System for Scientific Community. J. UCS 2011, 17, 1529–1549.
2. Alarfaj, F.; Kruschwitz, U.; Hunter, D.; Fox, C. Finding the Right Supervisor: Expert-Finding in a University Domain. In Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop, Montréal, QC, Canada, 3–8 June 2012; pp. 1–6.
3. Anderson, P. Perspective: Complexity Theory and Organization Science. Organ. Sci. 1999, 10, 216–232.
4. Balaid A.; Rozan M.; Hikmi S..; Memon J. Knowledge maps: A systematic literature review and directions for future research. International Journal of Information Management. 2016, 36, 451-475.
5. Balog, K.; Azzopardi, L.; De Rijke, M. A Language Modeling Framework for Expert Finding. Inf. Process. Manag. 2009, 45, 1–19.
6. Balog, K.; Azzopardi, L.; De Rijke, M. Formal Models for Expert Finding in Enterprise Corpora. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA, USA, 6–11 August 2006; pp. 43–50.
7. Balog, K.; Fang, Y.; De Rijke, M.; Serdyukov, P.; Si, L. Expertise Retrieval. Found. Trends® Inf. Retr. 2012, 6, 127–256.
8. Bargent, J. 11 Steps to Building a Knowledge Map. 2002
9. Berendsen, R.; Rijke, M.; Balog, K.; Bogers, T.; Bosch, A. On the Assessment of Expertise Profiles. J. Assoc. Inf. Sci. Technol. 2013, 64, 2024–2044.
10. Chan, K.; Liebowitz, J. The synergy of social network analysis and knowledge mapping: a case study. International Jounal of Management and Decision Making. 2006, 7(1), 19–35.
11. Chen, H.-H.; Treeratpituk, P.; Mitra, P.; Giles, C.L. Csseer: An Expert Recommendation System Based on Citeseerx. In Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries, Indianapolis, IN, USA, 22–26 July 2013; pp. 381–382.
12. Chung, W.; Chen, H.; Nunamaker Jr. J.F. Business intelligence explorer: a knowledge map framework for discovering business intelligence on the web. 2003. 36th Hawaii international conference on system sciences, pp. 1–10
13. Cifariello, P.; Ferragina, P.; Ponza, M. Wiser: A Semantic Approach for Expert Finding in Academia Based on Entity Linking. Inf. Syst. 2019, 82, 1–16.
14. Craswell, N.; Hawking, D.; Vercoustre, A.-M.; Wilkins, P. P@ Noptic Expert: Searching for Experts Not Just for Documents. In Ausweb Poster Proceedings; Southern Cross University: Queensland, Australia, 2001; p. 17.
15. Dang, Y.; Zhang, Y.; Chen, H.; Larson, C. C. Castillo-Chavez, H. Chen, W. B. Lober, M. Thurmond, & D. Zeng (Eds.), Knowledge mapping for bioterrorism-related literature. 2011, ,27, pp. 311–338
16. Das, S.; Mitra, P.; Giles, C. Learning To Rank Homepages for Researchername Queries. In the International Workshop on Entity-Oriented Search; ACM, EOS, SIGIR 2011 Workshop: Beijing, China, 2011; pp. 53–58.
17. Datta, A.; Yong, J.T.T.; Braghin, S. The Zen of Multidisciplinary Team Recommendation. J. Assoc. Inf. Sci. Technol. 2014, 65, 2518–2533.
18. Davoodi, E.; Afsharchi, M.; Kianmehr, K. A Social Network-Based Approach to Expert Recommendation System. In International Conference on Hybrid Artificial Intelligence Systems; Springer: Berlin/Heidelberg, Germany, 2012; pp. 91–102.
19. Davoodi, E.; Kianmehr, K.; Afsharchi, M. A Semantic Social Network-Based Expert Recommender System. Appl. Intell. 2013, 39, 1–13.
20. Deng, H.; Han, J.; Lyu, M.R.; King, I. Modeling and Exploiting Heterogeneous Bibliographic Networks for Expertise Ranking. In Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, Washington, DC, USA, 10–14 June 2012; pp. 71–80.
21. Deng, H.; King, I.; Lyu, M.R. Enhanced Models for Expertise Retrieval Using Community-Aware Strategies. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) 2012, 42, 93–106.
22. Driessen, S.; Huijsen, W.-O.; Grootveld, M. A framework for evaluating knowledge-mapping tools. Journal of Knowledge Management. 2007, 11(2), 109–117
23. Duffy, J. The KM technology infrastructure. Information Management Journal. 2000, 34(2), 62–66
24. Earl, M. Knowledge management strategies: toward taxonomy. Journal of Management Information Systems. 2001, 18(1), 215–233.
25. Ehrlich, K. Locating Expertise: Design Issues for an Expertise Locator System. Shar. Expertise-Beyond Knowledge Management; Mit Press: Cambridge, MA, USA, 2003; pp. 137–158.
26. Eppler, M. J. Making knowledge visible through knowledge maps: concepts, elements cases. Handbook on knowledge management (Vol. 1) Berlin Heidelberg: Springer. 2004, pp. 189–205
27. Eppler, M. J.; Simon, H. A. A process-based classification of knowledge maps and application examples. Knowledge and Process Management. 2008, 15(1), 59–71
28. Faisal, Hafiz & Rahman, Atta & Zaman, Gohar. Knowledge Mapping for Research Papers. 2019, 19, 158-164
29. Fang, Y.; Si, L.; Mathur, A. Facfinder: Search for Expertise in Academic Institutions; Technical Report: Serc-Tr-294; Department of Computer Science, Purdue University: West Lafayette, Indiana, 2008.
30. Fang, Y.; Si, L.; Mathur, A.P. Discriminative Graphical Models for Faculty Homepage Discovery. Inf. Retr. 2010, 13, 618–635.
31. Fazel-Zarandi, M.; Devlin, H.J.; Huang, Y.; Contractor, N. Expert Recommendation Based on Social Drivers, Social Network Analysis, and Semantic Data Representation. In Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems, Chicago, IL, USA, 27 October 2011; pp. 41–48. Appl. Sci. 2019, 9, 4250 29 of 32
32. Gilbert, J.; Raub, K. Managing knowledge-Building blocks for success, John Wiley & Sons. 2000, p. 360
33. Gollapalli, S.D.; Mitra, P.; Giles, C.L. Similar Researcher Search in Academic Environments. In Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, Washington, DC, USA, 10–14 June 2012; pp. 167–170.
34. Hansen, B. H.; Kautz, K. Knowledge mapping: a technique for identifying knowledge flows in software organisations. In Software process improvement. 2004. pp. 126–137. Berlin Heidelberg: Springer-Verlag.
35. Heck, T.; Hanraths, O.; Stock, W.G. Expert Recommendation for Knowledge Management in Academia. In Proceedings of the Association for Information Science and Technology, ASIST 2011, New Orleans, LA, USA, 9–13 October 2011; pp. 1–4.
36. Hellström, T.; Husted, K. Mapping knowledge and intellectual capital in academic environments: a focus group study. Journal of Intellectual Capital. 2004, 5(1), 165–180.
37. Hofmann, K.; Balog, K.; Bogers, T.; De Rijke, M. Contextual Factors for Finding Similar Experts. J. Assoc. Inf. Sci. Technol. 2010, 61, 994–1014.
38. Huang, C.; Yao, L.; Wang, X.; Benatallah, B.; Sheng, Q.Z. Expert as a Service: Software Expert Recommendation Via Knowledge Domain Embeddings in Stack Overflow. In Proceedings of the 2017 IEEE International Conference on Web Services (ICWS), Honolulu, HI, USA, 25–30 June 2017; pp. 317–324.
39. Huff A. Mapping strategic knowledge. London: Sage Publications. 2002, p. 302
40. Husain O.; Salim. N; Alias R.; Abdelsalam S.; Hassan A. Expert Finding Systems: A Systematic Review. Applied Sciences. 2019, 9, 4250
41. Juang, M.-C.; Huang, C.-C.; Huang, J.-L. Efficient Algorithms for Team formation with a Leader in Social Networks. J. Supercomput. 2013, 66, 721–737.
42. Kapur, R. Knowledge Management. 2020. Accessible online (last accessed 19.05.2020): https://www.researchgate.net/publication/338421805\_Knowledge\_Management
43. Kargar, M.; An, A. Discovering Top-K Teams of Experts with/without a Leader in Social Networks. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management, Glasgow, Scotland, UK, 24–28 October 2011; pp. 985–994.
44. Kim, S.; Suh, E.; Hwang, H. Building the knowledge map: an industrial case study. Journal of Knowledge Management. 2003, 7(2), 34–45
45. Kong, W.; Liu, Q.; Yang, Z.; Han, S. Collaborative Filtering Algorithm Incorporated with Cluster-Based Expert Selection. J. Inf. Comput. Sci. 2012, 9, 3421–3429.
46. Lecocq, R. Knowledge mapping: a conceptual model. 2006, pp. 1–96
47. Lee J.; Fink, D. Knowledge mapping: encouragements and impediments to adoption. Journal of Knowledge Management. 2013, 17(1), 16–28.
48. Li, C.-T.; Shan, M.-K.; Lin, S.-D. Context-Based People Search in Labeled Social Networks. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management, Glasgow, Scotland, UK, 24–28 October 2011; pp. 1607–1612.
49. Li, M.; Liu, L.; Li, C.-B. An Approach to Expert Recommendation Based on Fuzzy Linguistic Method and Fuzzy Text Classification in Knowledge Management Systems. Expert Syst. Appl. 2011, 38, 8586–8596.
50. Liebowitz, J. Linking social network analysis with the analytic hierarchy process for knowledge mapping in organizations. Journal of Knowledge Management. 2005, 9(1), 76–86.
51. Lin, S.; Hong, W.; Wang, D.; Li, T. A Survey on Expert Finding Techniques. J. Intell. Inf. Syst. 2017, 49, 1–25.
52. Liu, D.; Xu, W.; Du, W.; Wang, F. How to Choose Appropriate Experts for Peer Review: An Intelligent Recommendation Method in a Big Data Context. Data Sci. J. 2015, 14, 16.
53. Liu, D.-R.; Chen, Y.-H.; Kao, W.-C.; Wang, H.-W. Integrating Expert Profile, Reputation and Link Analysis for Expert Finding in Question-Answering Websites. Inf. Process. Manag. 2013, 49, 312–329.
54. Liu, L.; Li, J.; Lv, C. A method for enterprise knowledge map construction based on social classification. Systems Research and Behavioral Science. 2009, 205(2), 541–551
55. Logan, D.; Caldwell, F. Knowledge mapping: five key dimensions to consider. Gartner-Group. 2000, 30–50
56. Mangaravite, V.; Santos, R.L.; Ribeiro, I.S.; Gonçalves, M.A.; Laender, A.H. The Lexr Collection for Expertise Retrieval in Academia. In Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, Pisa, Italy, 17–21 July 2016; pp. 721–724.
57. Mansingh, G.; Osei-Bryson, K.-M.; Hirata, T. Building ontology-based knowledge maps to assist knowledge process outsourcing decisions. Knowledge Management Research & Practice. 2009, 7(1), 37–51
58. Mcdonald, D.W. Evaluating Expertise Recommendations. In Proceedings of the 2001 International ACM Siggroup Conference on Supporting Group Work, Boulder, CO, USA, 30 September–3 October 2001; pp. 214–223.
59. Mcdonald, D.W.; Ackerman, M.S. Expertise Recommender: A Flexible Recommendation System and Architecture. In Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA, USA, 2000; pp. 231–240.
60. Mockus, A.; Herbsleb, J.D. Expertise Browser: A Quantitative Approach To Identifying Expertise. In Proceedings of the 24th International Conference on Software Engineering, Orlando, FL, USA, 25–25 May 2002; pp. 503–512. Appl. Sci. 2019, 9, 4250 27 of 32
61. Moreira, C.; Wichert, A. Finding Academic Experts on A Multisensor Approach Using Shannon’s Entropy. Expert Syst. Appl. 2013, 40, 5740–5754.
62. Naeem, M.; Khan, M.B.; Afzal, M.T. Expert Discovery: A Web Mining Approach. J. Data Min. 2013, 1, 35–47.
63. Neshati, M.; Beigy, H.; Hiemstra, D. Expert Group Formation Using Facility Location Analysis. Inf. Process. Manag. 2014, 50, 361–383.
64. Neshati, M.; Beigy, H.; Hiemstra, D. Multi-Aspect Group formation Using Facility Location Analysis. In Proceedings of the 17th Australasian Document Computing Symposium, Dunedin, New Zealand, 5–6 December 2012; pp. 62–71.
65. Neshati, M.; Fallahnejad, Z.; Beigy, H. On Dynamicity of Expert Finding in Community Question Answering. Inf. Process. Manag. 2017, 53, 1026–1042.
66. Neshati, M.; Hiemstra, D.; Asgari, E.; Beigy, H. Integration of Scientific and Social Networks. World Wide Web 2014, 17, 1051–1079.
67. Pei, X.; Wang, C. A study on the construction of knowledge map in matrix organizations. International conference of knowledge management and information retrieval, 2009. (pp. 102–107)
68. Petkova, D.; Croft, W.B. Proximity-Based Document Representation for Named Entity Retrieval. In Proceedings of the 16th ACM Conference on Conference on Information and Knowledge Management, Lisbon, Portugal, 6–10 November 2007; pp. 731–740.
69. Price, S.; Flach, P.A. Computational Support for Academic Peer Review: A Perspective From Artificial Intelligence. Commun. ACM 2017, 60, 70–79.
70. Protasiewicz, J.; Pedrycz, W.; Kozłowski, M.; Dadas, S.; Stanisławek, T.; Kopacz, A.; Gał ˛e˙zewska, M. A Recommender System of Reviewers and Experts in Reviewing Problems. Knowl. Based Syst. 2016, 106, 164–178.
71. Quintas, P. Implications of the division of knowledge for innovation in networks. In Networks, alliances and partnerships in the innovation process. Springer US. 2002, pp. 135–162.
72. Rao, L.; Mansingh, G.; Osei-Bryson, K.-M. Building ontology based knowledge maps to assist business process re-engineering. Decision Support Systems. 2012, 52(3), 577–589
73. Renukappa, Egbu C. Knowledge mapping: concepts and benefits for a sustainable urban environment. The 20th annual conference association of researchers in construction management (ARCOM). 2004, Vol. 2 (pp. 1–12).
74. Ribeiro, I.S.; Santos, R.L.; Gonçalves, M.A.; Laender, A.H. On Tag Recommendation for Expertise Profiling: a Case Study in the Scientific Domain. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 189–198.
75. Sateli, B.; Löffler, F.; König-Ries, B.; Witte, R. Scholarlens: Extracting Competences from Research Publications for the Automatic Generation of Semantic User Profiles. PeerJ Comput. Sci. 2017, 3, 121.
76. Silva, A.T.P. A Research Analytics Framework for Expert Recommendation in Research Social Networks; City University of Hong Kong: Hong Kong, China, 2014.
77. Silva, T.; Guo, Z.; Ma, J.; Jiang, H.; Chen, H. A Social Network-Empowered Research Analytics Framework for Project Selection. Decis. Support Syst. 2013, 55, 957–968.
78. Smirnova, E.; Balog, K. A User-Oriented Model for Expert Finding. Adv. Inf. Retr. 2011, 6611, 580–592.
79. Sriharee, G.; Anekboon, K. Multiple Aspect Ranking for Researcher Expertise Finding. In Proceedings of the IEEE International Conference on Knowledge Engineering and Applications (Ickea), Singapore, 28–30 September 2016; pp. 207–211.
80. Stankovic, M.; Wagner, C.; Jovanovic, J.; Laublet, P. Looking for Experts? What Can Linked Data do for You? Available online: https://www.researchgate.net/publication/221023369\_Looking\_for\_Experts\_What\_can\_Linked\_Data\_do\_for\_You (accessed on 10 May 2020).
81. Sun, J.; Xu, W.; Ma, J.; Sun, J. Leverage Raf to Find Domain Experts on Research Social Network Services: A Big Data Analytics Methodology With Mapreduce Framework. Int. J. Prod. Econ. 2015, 165, 185–193.
82. Sziklai, B. How to Identify Experts in a Community? Int. J. Game Theory 2018, 47, 155–173.
83. Tang, J.; Zhang, J.; Yao, L.; Li, J.; Zhang, L.; Su, Z. Arnetminer: Extraction and Mining of Academic Social Networks. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NV, USA, 24–27 August 2008; pp. 990–998.
84. Taylor, M.; Richards, D. Finding and Validating Expertise; ECIS: London, UK, 2011.
85. Tiwana, A. Knowledge management toolkit, the amrit tiwana knowledge management toolkit. First Edit.1999, pp. 1–482.
86. Torkzadeh Mahani, N.; Dehghani, M.; Mirian, M.S.; Shakery, A.; Taheri, K. Expert Finding by the Dempster-Shafer Theory for Evidence Combination. Expert Syst. 2018, 35, 12231.
87. Tran, H.D.; Cabanac, G.; Hubert, G. Expert Suggestion for Conference Program Committees. In Proceedings of the 11th International Conference on Research Challenges in Information Science (Rcis), Brighton, UK, 10–12 May 2017; pp. 221–232.
88. Uddin, M.N.; Duong, T.H.; Oh, K.-J.; Jo, G.-S. An Ontology Based Model for Experts Search and Ranking. In Asian Conference on Intelligent Information and Database Systems; Springer: Berlin/Heidelberg, Germany; pp. 150–160.
89. Vestal, W. Knowledge mapping: the essentials for success. Houston, TX: APQC Publications. 2005 p. 75.
90. Wang, G.A.; Jiao, J.; Abrahams, A.S.; Fan, W.; Zhang, Z. Expertrank: A Topic-Aware Expert Finding Algorithm for Online Knowledge Communities. Decis. Support Syst. 2013, 54, 1442–1451.
91. Wang, Q.; Ma, J.; Liao, X.; Du, W. A Context-Aware Researcher Recommendation System for Universityindustry Collaboration on R&D Projects. Decis. Support Syst. 2017, 103, 46–57.
92. Weiss, D.J.; Shanteau, J. Empirical Assessment of Expertise. Hum. Factors 2003, 45, 104–116.
93. Wexler, M. N. The who, what and why of knowledge mapping. Journal of Knowledge Management. 2001, 5(3), 249–264
94. Wu, C.-J.; Chung, J.-M.; Lu, C.-Y.; Lee, H.-M.; Ho, J.-M. Using Web-Mining for Academic Measurement and Scholar Recommendation in Expert Finding System. In Proceedings of the 2011 IEEE/Wic/Acm International Conference on Web Intelligence and Intelligent Agent Technology (Wi-Iat), Lyon, France, 22–27 August 2011; pp. 288–291.
95. Wu, H.; He, J.; Pei, Y.; Long, X. Finding Research Community in Collaboration Network with Expertise Profiling. In International Conference on Intelligent Computing; Springer: Berlin/Heidelberg, Germany, 2010; pp. 337–344.
96. Xu, W.; Sun, J.; Ma, J.; Du, W. A Personalized Information Recommendation System for R&D Project Opportunity Finding in Big Data Contexts. J. Netw. Comput. Appl. 2016, 59, 362–369.
97. Xu, Y.; Guo, X.; Hao, J.; Ma, J.; Lau, R.Y.K.; Xu, W. Combining Social Network and Semantic Concept Analysis for Personalized Academic Researcher Recommendation. Decis. Support Syst. 2012, 54, 564–573.
98. Yang, M.-S.; Kang, N.-K.; Kim, Y.-J.; Kim, J.-S.; Choi, K.-N.; Kim, Y.-K. Expert Recommendation System Based on Analyzing Expertise and Networks of Human Resources in National Science & Technology Information Service. J. Central South Univ. 2013, 20, 2212–2218.
99. Yasin, F.; Egbu, C. Critical steps to knowledge mapping in facilities management organisation. 27th annual ARCOM conference. 2011, 603–612.
100. Yimam-Seid, D.; Kobsa, A. Expert-Finding Systems for Organizations: Problem and Domain Analysis and the Demoir Approach. J. Organ. Comput. Electron. Commer. 2003, 13, 1–24.
101. Zhan, Z.; Yang, L.; Bao, S.; Han, D.; Su, Z.; Yu, Y. Finding Appropriate Experts for Collaboration. In International Conference on Web-Age Information Management; Springer: Berlin/Heidelberg, Germany, 2011; pp. 327–339.
102. Zhao, G.; Miao, P.; Guan, Y. Construction of intelligence knowledge map for complex product development. Journal of Engineering Science and Technology Review. 2013, 6(3), 82–87

# Appendix 1

General questions

1. Expertise is most often referred to as the human cognitive skills that can be developed by conducting a task frequently. Do you have any objections to such definition?
2. Do you consider expertise location effective if scoped only within university or in order to be considered useful it has to include several organizations?
3. What do you think about opportunity for experts to manually correct personal profiles? Do you feel there are additional risks implied by expansion of the model to several universities? Does potentially increased precision outweigh those risks?

Questions for academic directors

1. Which are the main reasons for you to locate experts: understanding overall organization expertise to understand capabilities for creating new and developing existing courses, assigning professor that fits particular course best, understanding gaps in overall organization knowledge?
2. Which sources would you use to verify expertise in particular domain for each of those purposes: courses taught, courses developed, publications within this course scope, professional experience in this area, h-index, conference participation, something else?
3. What are important requirements or aspirations for expert location process for each of the purposes: possibility of comparison with external experts and organizations, location of not only existing expertise but also attempt to locate expertise missing in the organization to identify knowledge gaps, data relevance, data precision; something else?
4. What are your main concerns about possible problems?

Questions for researchers

1. Which are the main reasons for you to locate experts: finding collaborator for conducting particular research, finding expert for paper reviewing, something else?
2. Which sources would you use to verify expertise in particular domain for each of those purposes: publications in the same or related domain, h-index, methodology used in previous research, participation in conferences, something else?
3. What are important requirements or aspirations for expert location process for each of the purposes: high level of detalization of expertise to choose collaborator whose expertise covers most relevant domains required for this research that searching researcher lack, personal fit with researcher looking for collaborator: age, overlapping or complementary expertise, something else?
4. What are your main concerns about possible problems?

Questions for students

1. Which are the main reasons for you to locate experts, besides finding academic supervisor for writing thesis?
2. Which sources would you use to verify expertise in particular domain for this purpose(s): previously supervise theses, h-index, preferred methodology both within own research and supervised theses?
3. What are important requirements or aspirations for expert location process for each of the purposes: common research interests, preferred methodology match: qualitative or quantitative, system ability to advise most suitable domains, based on student personal interests, as well as available supervisors’ ones, integration with the official system of assigning supervisors, something else?
4. What are your main concerns about possible problems?

Questions for experts

1. Since you clearly possess expertise and so far are a subject for expertise location process, what is the most important for you in this process: precision of the results, required amount of self-disclosed data, data security, time resources required, something else?

# Appendix 2

Domain structure

**Finance and accounting:** audit, business planning, internal control, investment, quantitative research methods in finance, corporate governance, international financial management, emerging markets, risk management, strategic finance, intellectual capital management, management accounting, financial modeling, financial planning and forecasting, financial institutions and markets, financial analysis, financial accounting, fintech, value-oriented management.

**International management and international business:** global trade, global economy, global economy, international institutional environment, international trade, international entrepreneurship, international business strategies, international institutions, international supply chain operations, international marketing, international financial management, multinational companies (MNCs), foreign direct investment (FDI), emerging markets, strategies for entering foreign markets.

**Information systems and information management in management**: enterprise architecture, knowledge engineering, artificial intelligence, IT in logistics and supply chain management, IT in HR management, corporate information systems, emerging markets, knowledge management systems, smart cities, business management processes, data management, knowledge management, sustainability, fintech, digital commerce, digital transformation, digital marketing, e-government and IT in state and municipal government.

**Entrepreneurship and innovation**: business planning, venture financing, corporate entrepreneurship, international entrepreneurship, youth entrepreneurship, new entrepreneurial teams, open innovation, entrepreneurship, emerging markets, the development of small and medium-sized businesses, family business, the creation of new firms, social entrepreneurship, strategic entrepreneurship, technological innovation, innovation management, sustainable innovation, effectuation and causation.

**Marketing and sales**: B2B marketing, innovation marketing, relationship marketing, marketing research, marketing channels, marketing communications, international marketing, consumer behavior, emerging markets, market orientation, strategic marketing, brand management, sales management, pricing, digital commerce, digital marketing.

**Organizational behavior and human resource management**: coaching, IT in HR management, corporate governance, leadership, staff motivation and compensation, staff assessment and hiring, success psychology, emerging markets, human resource development, labor markets, professional and personal development technologies, management knowledge, team management, conflict management, HRM in innovative companies, economic psychology.

**Economics, ecosystems, markets and institutions**: global economy, institutional economics, macroeconomics, macroeconomic policies in the global environment, international political economics, microeconomics, emerging markets, oil and gas markets, labor markets, econometrics, economics for managers, energy markets.

**Methods of data analysis and decision making**: business statistics, simulation, operations research, qualitative methods of data analysis and decision making, quantitative methods of data analysis and decision making, mathematical methods in economics, research methods, emerging markets, statistics, game theory, econometrics.

**State and municipal governance (SMG) and urban studies**: anti-corruption policy, urban transport policy, public procurement, public service, public-private partnerships, institutional analysis of public policy, macroeconomic modeling and forecasting, taxation and insurance, emerging markets, smart cities, management of public sector organizations, management territorial development, sustainability, health economics, economics and management of infrastructure sectors (natural monopolies), social economics, e-government and IT in the SMG.

**Strategic management**: organizational development management, enterprise architecture, business planning, corporate social responsibility, corporate entrepreneurship, corporate governance, international business strategies, responsible leadership, emerging markets, strategic orientation, strategic analysis, strategic leadership, strategic technological innovation management, management knowledge, change management, intellectual capital management, enterprise transformation management, performance management, sustainability, value-oriented management, digital transformation.

**Operations and project management**: logistics, analytical tools in supply chain management, business planning, IT in logistics and supply chain management, international operations in the supply chain, emerging markets, distribution in the supply chain, warehouse logistics, transport logistics, business management processes, inventory management, knowledge management, quality management, project management, supply chain management, performance management, sustainability, sustainable value chains.

**Interdisciplinary and other areas**: soft skills, enterprise architecture, business ethics, visual and analytical thinking, business communications, design thinking, knowledge engineering, case study method, teaching business disciplines, research methods, negotiations, rhetoric, knowledge management, intellectual capital management.

# Appendix 3

Expert profile visualization

This is profile constructed for the query, expressing need in research supervisor and having preset proportions 40%–40%–20%. Domain 1 has 70% of importance, while Domain 2 only 30%.

Expert ranked list

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Domain 1** | | | | **Domain 2** | | | | **Overall** |
|  | Research | Teaching | Applied | **Overall** | Research | Teaching | Applied | **Overall** |
| **Expert 1** | 0.9 | 0.8 | 0.5 | **0.78** | 0.6 | 0.5 | 0.5 | **0.54** | **0.708** |
| **Expert 2** | 0.7 | 0.8 | 0.6 | **0.72** | 0.5 | 0.9 | 0.4 | **0.64** | **0.696** |
| **Expert 3** | 0.8 | 0.6 | 0.6 | **0.68** | 0.5 | 0.6 | 0.2 | **0.48** | **0.62** |
| **Expert 4** | 0.5 | 0.4 | 0.7 | **0.5** | 0.7 | 0.6 | 0.4 | **0.6** | **0.53** |
| **Expert 5** | 0.3 | 0.7 | 0.2 | **0.44** | 0.1 | 0.2 | 0.1 | **0.14** | **0.35** |

# Appendix 4.

Overall organization expertise knowledge map

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | |  |  |  |
| **Expert 1** | **Expert 2** | **Expert 3** |
|  | **Domain 1** |  |  |  |
|  | Domain 1.1 |  |  |  |
|  | Domain 1.2 |  |  |  |
|  | **Domain 2** |  |  |  |
|  | **Domain 3** |  |  |  |