University of Magdeburg

School of Computer Science

Master's Thesis

Supporting Quality Assessment in Systematic Literature Reviews

Author:

Yusra Shakeel

October 23, 2017

Advisors:

Prof. Dr. rer. nat. habil. Gunter Saake,
Prof. Dr.-Ing. Thomas Leich,
Dipl.-Inform. Ivonne von Nostitz-Wallwitz,
M.Sc. Jacob Krüger

Otto-von-Guericke-University Magdeburg

Shakeel, Yusra:
Supporting Quality Assessment in Systematic Literature Reviews
Master's Thesis, University of Magdeburg, 2017.

Abstract

The amount of empirical studies published in the field of computer science is rapidly increasing, resulting in an increase in the number of literature reviews. Researchers have been conducting systematic literature reviews to summarize the existing evidence regarding a research area or topic of interest, identify gaps in the current research and build new research activities. Despite of their usefulness, executing a systematic literature review requires considerable amount of time and effort. Thus, apart from conducting, researchers are constantly proposing semi-automated approaches to support different stages of the review process. The most laborious and error-prone step during execution of a systematic literature review process is the selection of primary studies. Therefore, within the scope of present work, a systematic literature review is conducted to analyze existing strategies proposed by software engineering researchers to support the selection of primary studies. Based on the findings, a semi-automated approach is presented to assist reviewers with this essential step during the review process. The strategy is implemented prototypically to examine the validity based on performance measures, such as, accuracy, precision and recall. The results show that through the proposed approach, selection and quality assessment of primary studies during the review process can be partially automated. However, there are certain limitations for the approach that must be further investigated to obtain more promising results.

Acknowledgement

First of all, I would like to thank Prof. Gunter Saake for providing me with the opportunity to work on my master thesis within his working group.

It gives me great pleasure in acknowledging the support of Prof. Thomas Leich throughout. His positive attitude towards every matter kept me motivated during my entire work duration at the METOP GmbH.

I would also like to express my sincere gratitude to Ivonne von Nostitz-Wallwitz and Jacob Krüger for their continuous support and encouragement throughout my work. In spite of being surrounded by other work commitments, they were always there to guide me. This master thesis could not have been possible without their excellent supervision, regular discussions and constructive feedbacks.

Many thanks to everyone at the METOP GmbH for providing me a great working atmosphere and experience. Finally, I must express my very profound gratitude to my parents, Shakeel-ur-Rehman and Zeenat Shakeel for their immense support and prayers throughout my years of study. Last but not least, I thank my husband, Jawwad for continuously encouraging me to achieve my goals.

Contents

List of Figures

List of Tables

List of Acronyms

CMM	Capability Maturity Model
CS	Computer Science
EC	Exclusion Criteria
GSD	Global Software Development
IC	Inclusion Criteria
IR	Information Retrieval
JSON	JavaScript Object Notation
KDD	Knowledge Discovery In Databases
PDF	Portable Document Format
PEx	Projection Explorer
RQ	Research Question
SCAS	Score Citation Automatic Selection
SE	Software Engineering
SLR	Systematic Literature Review
VTM	Visual Text Mining

1. Introduction

Originating from the medical domain, systematic literature reviews (SLRs) became an important research method in software engineering (SE) to summarize existing evidence [Babar and Zhang, 2009; Kitchenham et al., 2004; Webster and Watson, 2002]. SLRs provide information to help researchers identify areas for improvement in the current research and appropriately position new research activities [Kitchenham and Charters, 2007]. As a result, the number of SLRs have rapidly increased over the last years [da Silva et al., 2011; Kitchenham et al., 2009; Verner et al., 2012]. Due to their usefulness and significant features, SLRs have been adopted by many other domains, for instance software engineering, criminology, social policy, economics, infrastructure and nursing [Kitchenham et al., 2004, 2009].

The concept of evidence-based software engineering was introduced as a useful research approach by Kitchenham et al. [2004]. Since their first introduction in 2004, SLRs have gained significant attention from researchers of the SE community. The number of SLRs reported in the field of computer science (CS) further increased after the first adapted guidelines for undertaking systematic reviews in SE were published by Kitchenham and Charters [2007]. The volume of research that needs to be considered by SE researchers is constantly expanding, making it difficult to manually read, critically evaluate, and synthesize the state of current knowledge. Hence, SE researchers have been conducting and reporting more and more SLRs on diverse topics, such as, agile software development, regression testing, process modeling, variability management, cost estimation, organizational motivators for Capability Maturity Model (CMM)-based process improvement, and statistical power [Babar and Zhang, 2009]. Still, despite their value in SE, conducting SLRs remains a time consuming and difficult task [Felizardo et al., 2011; Hassler et al., 2016]. In particular, tools to conduct SLRs more efficiently and semi-automatic may support different steps but are often limited or even missing.

1.1 Motivation

The number of empirical studies reported in SE have significantly increased since 2004, but there are some problems associated with them. The implementation of SLRs encounter difficulties due to shortcomings in reporting empirical studies, for example, the approach used to conduct the study is not clearly explained, conclusions are incomplete or questionable [Brereton et al., 2007]. Hence, SE researchers, apart from conducting and reporting SLRs, also focus on improving the SLR methodology and quality. For example, providing techniques to assess the quality of the reported primary studies included in an SLR [Zhang and Babar, 2013]. Although SLRs have proved to be an important research methodology for evidence-based software engineers, the SLR process can be very time consuming and difficult [Felizardo et al., 2011].

To perform the SLR process more efficiently and effectively, SE researchers are proposing semi-automated approaches to assist various aspects of an SLR. Adequate tool support for different stages reduce the effort and time required to conduct an SLR [Hassler et al., 2016]. The review process usually involves a large volume of potentially relevant literature regarding the topic of interest. Hence, one of the most laborious step while conducting an SLR is to identify and select suitable primary studies [Felizardo et al., 2011; Schröter et al., 2017a]. In their study, Hassler et al. [2016] also identified that tool support for this task is one of the most desired ones by the community. Thus, it would be advantageous for researchers to design approaches that accelerate the execution phase of an SLR [Molleri and Benitti, 2012].

The conducting phase of an SLR aims at retrieving relevant primary studies with minimum bias and maximum validity. The quality of a literature review highly depends on the quality of the included primary studies hence, quality assessment is an important SLR activity [Brereton et al., 2007; Dybå and Dingsøyr, 2008a]. In most cases, quality assessment checklists defined by the reviewers are used to manually determine quality of primary studies but that can be tedious and challenging [Kitchenham and Charters, 2007]. Therefore, any method to support reviewers with quality assessment of primary studies will be very useful [da Silva et al., 2011; Hassler et al., 2016]. The existing approaches, discussed later in Section 3.3, focus on obtaining articles based on their content and citation relationships. However, to ensure reliable SLR results, the selected relevant studies further need to be critically evaluated for validity. In this work, we extend the previous approaches and aim to obtain the most promising primary studies by scoring them based on defined quality criteria.

1.2 Goal of this Thesis

The challenges we describe above have an overall effect on the efficiency and validity of an SLR. The set of possibly relevant articles obtained after the search process must be assessed to identify those that satisfy the defined quality criteria. Therefore, the aim of this thesis is to develop an approach that will support evaluation of primary studies based on their content and quality. For this purpose, we analyze the existing

approaches that support the selection of primary studies in an SLR. We use the results to derive an approach that effectively performs the selection and quality assessment activity during an SLR process. Furthermore, we evaluate the proposed approach to investigate validity of the results.

In this work, we focus on reducing the time and effort required for an SLR by suggesting an approach that supports the selection activity. To achieve the goals of this thesis we work through the following steps:

- We conduct a literature review to identify and synthesize the current state-of-the-art on semi-automatic approaches for the selection of primary studies for SLRs. The aim is to study the methodologies proposed previously and identify areas that demand further research.
- Based on the findings, we propose an approach to retrieve relevant articles using specific parts of the content (title, abstract, keywords) and citation relationships. Furthermore, we assign a score to each article based on the defined quality criteria.
- Finally, we evaluate the proposed approach by using an example case study. We examine the primary studies selected by the example SLR conducted with the results we obtain through our approach. To investigate validity of the results and flexibility of the procedure, we determine performance measures, such as, accuracy, precision and recall.

Implementing these procedures will greatly support the SLR process aiding researchers to efficiently identify most relevant studies related to their topic of interest. Quality assessment along with content and reference relationship combined, will result in more reliable results obtained by conducting an SLR.

1.3 Structure of the Thesis

In **Chapter 2** we introduce the general concepts and background information regarding SLRs, text mining and information retrieval. We aim to describe the basis of our work in this chapter to assist readers unfamiliar with the topic.

We begin **Chapter 3** with a description of an SLR we conduct to analyze the existing approaches that support selection of primary studies for CS researchers. Further in this chapter we explain the complete research method for execution of the review and analysis of the results we obtain.

In **Chapter 4** we propose an approach to support efficient conduct of an SLR. We derive the approach using results obtained from the SLR elaborated in previous chapter. Later, we explain the methodology that we recommend in this thesis work to obtain the most promising studies relevant to a specific research topic or question.

Within **Chapter 5** we explain the implementation and evaluation of the proposed approach. We discuss the performance and validity of results by conducting an existing study using our approach.

In **Chapter 6** we discuss the conclusions and summary of the presented thesis. Furthermore, we explain the future work that is possible in this area of research.

2. Background

In this chapter we explain the background information regarding SLRs and data mining. We begin with the detailed description of the phases of an SLR process and the involved stages. Later on, we elaborate the concepts of text mining for a better understanding of the approaches to support an SLR that we discuss later in the next chapters. We discuss some specific text mining methods, such as, information extraction, automatic term recognition, document classification and clustering, along with the text analysis process.

2.1 Systematic Literature Review

An SLR is a research methodology used to identify, evaluate and interpret all the available studies relevant to a particular research question (RQ) , topic area or phenomenon of interest [Kitchenham, 2004]. The individual studies that contribute to a systematic review are called primary studies therefore, an SLR is a form of secondary study [Kitchenham, 2004]. An SLR summarizes the existing evidence regarding a research topic that helps identify any gaps for further improvement and provide a basis to appropriately position new research activities [Kitchenham and Charters, 2007].

To conduct research in any field of study, it is important to perform a literature review to build good understanding regarding the topic of interest. The main advantage of an SLR is that it follows a well-defined procedure to answer defined RQs, making the process transparent and repeatable up-to a certain degree [Boell and Cecez-Kecmanovic, 2015]. Another kind of literature review are the traditional narrative reviews that are informative but usually limited to a subset of studies in the chosen area, therefore can result in incomplete conclusions [Booth et al., 2012]. The SLR process differs from a conventional expert literature review in the following ways as indicated by Kitchenham and Charters [2007]:

- An SLR begins with defining a review protocol that specifies the RQs of interest and methods that will be used to perform the review.

- The search strategy is defined, to identify as much of the relevant literature as possible and ensure transparency of search process.
- To assess each of the potentially relevant primary studies, inclusion and exclusion criteria are stated.
- For evaluating each primary study, the quality criteria are specified.
- To extract the required information from relevant literature, data extraction forms are designed.

Therefore, the likelihood of selection bias is reduced by carefully defining the review protocol at the beginning of the SLR process [Booth et al., 2012]. The conclusions drawn from an SLR are more reliable and clear compared to a narrative review but successful completion of the SLR process can be challenging [Kitchenham and Charters, 2007]. Some level of expertise in the topic of interest is also advantageous for an effective SLR [Booth et al., 2012]. The RQs for the literature review should be focused on an issue which has not been addressed previously in order to avoid unnecessary effort. A good search strategy must be defined followed with an efficient study selection criteria to ensure valid results. The quality of an SLR highly depends on the quality of primary studies hence quality assessment must be performed for each of them [Dybår et al., 2007; Kitchenham and Charters, 2007].

The SLR methodology has been broadly adapted by the evidence based researchers of the SE community following the guidelines published by Kitchenham and Charters [2007]. The SLR process involves a number of discrete activities grouped into three main phases; planning the review, conducting the review, reporting the review. We summarize the three phases with their composite stages in Figure 2.1. Further in this section we provide a detailed explanation about each of the stages as published by Kitchenham and Charters [2007].

2.1.1 Phase 1: Planning

An SLR begins with the planning phase during which purpose of conducting research on a specific topic is defined. The review protocol that defines review procedures to be followed is carefully prepared as it is an essential part of the SLR process. In some situations, where the systematic reviews are commissioned, a commissioning document must also be prepared during this phase. The most significant activity involved in the planning phase is defining the RQs and preparing a review protocol. Before initiating the SLR process, reviewer must conduct a scoping review about the topic of interest to ensure if there is a need to address the particular issue. Furthermore, to conduct a secondary review it must be identified if there is sufficient information available related to the topic that can be used as primary studies. The process must not start if the RQ has been addressed previously or if there is not enough evidence available to support the research. As a result of planning phase, a review plan acting as a road map for the process is obtained which must be followed to conduct an SLR.

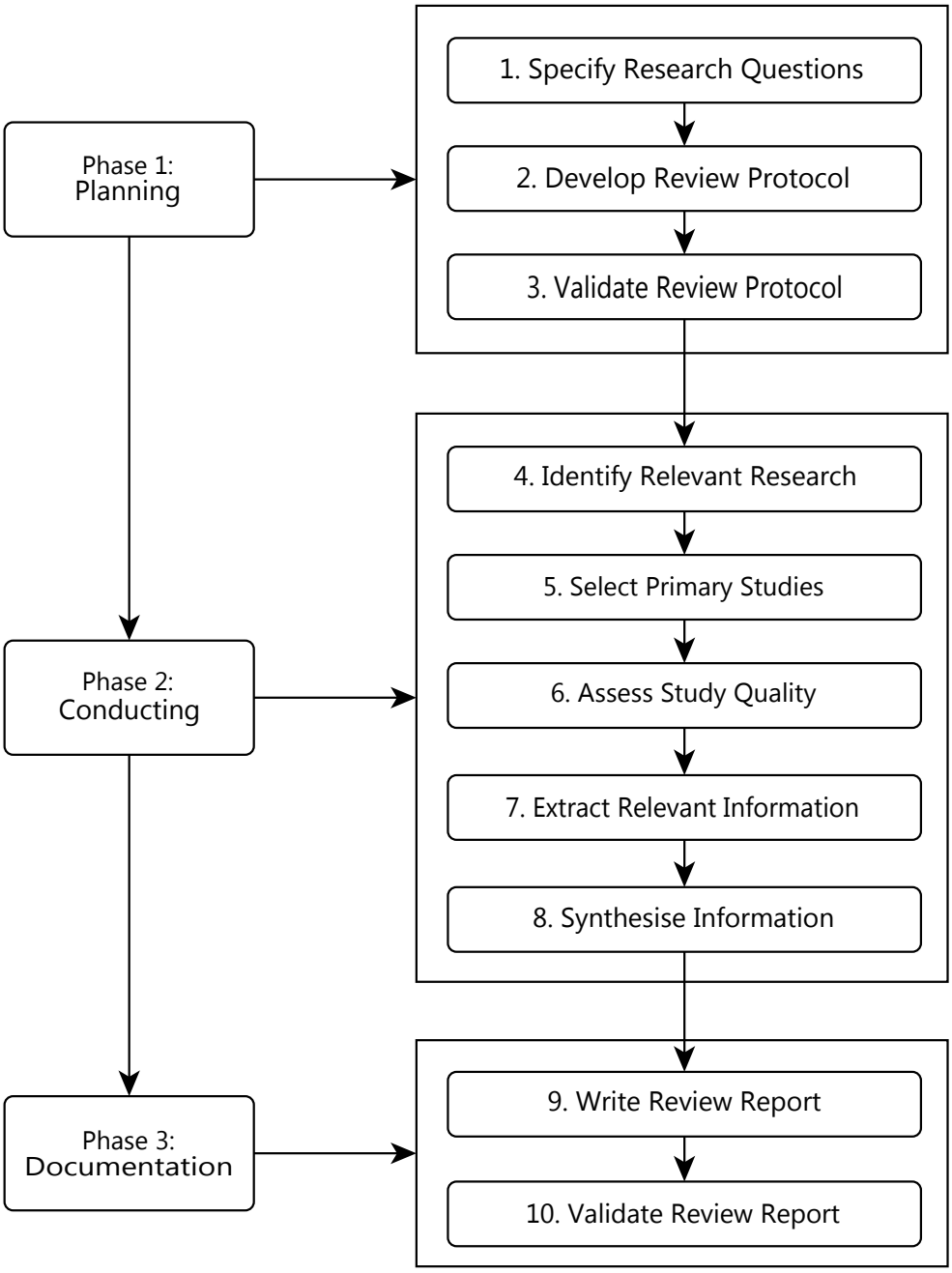

Figure 2.1: An SLR process, adapted from Kitchenham and Charters [2007].

Specify RQs

The most important step during planning phase of an SLR is defining the RQs to be answered. The RQs must be meaningful and interesting for the researchers, aiming to either identify existing studies or build future research activities. The review questions form the foundation for an SLR, guiding the review process to identify relevant primary

studies and analyze information extracted from them [Dybår et al., 2007]. To formulate the RQs in a way that is clear and precise, a PICOC (Population, Intervention, Comparison, Outcome, Context) criteria is recommended for reviewers, explained as:

- Population: A question must define the specific population group, such as, researchers working in an application area or a category of software engineers, that will benefit from the study being conducted.
- Intervention: The software methodology, tool, technology, or procedure that addresses a specific issue is defined as the intervention.
- Comparison: For reviewing the interventions, it is required to define appropriate comparison of interest.
- Outcome: This component describes the relevant and quantifiable characteristics important to the practitioners, such as, improved reliability, or reduced time.
- Context: This is the context in which the comparison is performed, for instance, academic or industrial, the participants of the study and the tasks being performed.

Develop a review protocol

Independent of the field of research, every SLR requires a detailed protocol to be prepared beforehand based on ideas from all members of the review team. The research protocol acts as a plan, describing the process and methods that will be applied throughout an SLR. It also contains some information regarding review team members and their roles during the process. The review team can also seek expert advice regarding the topic for a better understanding of their work. A pre-defined protocol reduces the possibility of researcher bias as it consists of all the information required to conduct next stages of an SLR. We discuss the components that must be part of review plan below:

Background and rationale: This section of the review protocol consists of a brief summary regarding current research about the topic of interest. Moreover, motivation of the reviewer to perform an SLR addressing the specific issue and the need for a solution must be stated.

RQs: RQs must be clearly defined in the plan, as it forms basis of the research being performed. The questions should be precise and focusing on a specific issue to limit scope of the review. PICOC criteria is recommended to structure the question(s), as any modification later can completely change purpose of the review being conducted.

Search strategy: The search method to retrieve primary studies including search terms and resources to be searched are defined in this component. RQs are used to extract keywords and generate a search query. It must include all possible synonyms of the keywords, their abbreviations and alternative spellings. Sophisticated search strings are then constructed using Boolean ANDs and ORs. The method that

will be used to search for relevant studies must be determined. The researchers should also justify the way specific searching methods, such as, manual searching, automated searching, snowballing and contacting key researchers are combined. For an automated search, the description of data resources and search query that will be used is specified in this section. If a manual search is planned, suitable journal and conference proceedings must be defined along with justification of the selection.

Study selection criteria: In this component of the review protocol, the researcher must specify the study selection criteria to determine whether a study should be included for an SLR or not along with the procedure that will be followed to apply the defined criteria. The inclusion criteria (IC) and exclusion criteria (EC) are formulated based on the defined RQs and searched studies need to be sorted in such a manner that only those fulfilling the IC are included as primary studies. The study selection criteria can be applied in different stages, initial decisions are usually based in specific parts of the paper, such as, title, abstract and keywords, to exclude the obvious irrelevant ones. Later on, the full-text of candidate papers are analyzed by the reviewers for a final decision regarding their inclusion. The restriction to the number of primary studies is specified to focus on the issue being addressed and make the SLR process practically manageable.

Quality assessment checklist and procedures: The review team must agree on quality assessment procedures to determine the quality of primary studies included in an SLR. The minimum acceptance criteria are specified in this component along with the procedure to assess all primary studies. The credibility of an SLR is assured by including valid and reliable primary studies. To perform quality assessment manually, reviewers must develop quality criteria checklists based on study type, for example, qualitative or quantitative study. Each article is scored according to their fulfillment of individual criteria, the total points are then assigned. The review team must also define methods to resolve conflicts between their decisions because the assessment is done manually by members, according to their individual expertise regarding the topic.

Data extraction strategy: After screening and quality testing of primary studies, data to be retrieved from the individual studies must be stated. The researchers must also define the procedures for performing the extraction and validation of the data. Reviewers are recommended to prepare data extraction forms for consistency in the extracted data. The format of these forms highly depend on the quantitative or qualitative nature of primary studies. Data to be stored for all the papers include their publication details along with information to be used for answering RQs. These forms are filled by review team members for each of the selected studies. Strategies to resolve conflicts between individual decisions must also be stated as part of this component.

Data synthesis and aggregation procedure: This part of the review protocol defines strategies to summarize, integrate, combine and compare findings from the primary studies selected for an SLR. For qualitative study type, synthesis is usually

an iterative process as authors might explain concepts differently from each other. The strategy to combine findings of multiple studies must be defined in a manner appropriate to answer the RQs.

Strategy for knowledge translation: It is important to note the target audience who would benefit from the results of SLR being conducted. The procedures used to spread the obtained information needs to be defined, usually researchers publish their findings as journal publications or conference articles.

Project timetable: The time frame for conducting an SLR must be decided according to the complexity of review. The work plan for each of the proposed steps must be determined to keep track of the process. To identify the time requirement for conducting an SLR, initial screening regarding the topic can be performed to determine its complexity.

Validate review protocol
Review protocol is one of the most crucial part of an SLR hence, it must be evaluated to verify its rigor. In this step, the reviewers specify procedures to be used for validating the protocol. Both, internal and external validation is recommended as a complete procedure. Review team members validate specific aspects of the protocol, such as, the search strategy, data extraction forms as well as data synthesis procedures. External validation involves evaluation of the review plan by experts working in the research area of interest.

2.1.2 Phase 2: Conducting

As a result of the planning phase, a well-prepared and verified review plan to conduct the SLR is obtained. Since the information required to perform stages of execution phase are part of the review protocol, reviewers must follow the plan to achieve results. The conducting phase involves identification of relevant research, selection and quality assessment of primary studies. Furthermore, data extraction and information synthesis are also part of this phase. The reviewer retrieves primary studies that answer the defined RQs as a result of this phase. We discuss each of the steps involved for executing an SLR further in this section.

Identify relevant research
The conducting phase of an SLR begins with applying the search strategy defined in review protocol. The goal of an SLR is to retrieve as many relevant primary studies as possible to answer the RQs. The search strategy for an SLR must be unbiased to find maximum relevant papers from the resources used. Usually researchers prefer to perform automated search using resources, such as, digital libraries and indexing systems. Other searching methods include manual searching of selected journals and conference proceedings as well as snowballing methods. To identify all the relevant papers both, forward and backward snowballing is performed. The search strategy must combine the searching methods in such a way that an acceptable level of completeness is achieved. The reviewers must document their searches as efficiently as possible to assure

the repeatability and transparency of an SLR. The resources used to find candidate papers must be selected prior to execution phase. These usually include digital libraries, journal and conference proceedings and other resources such as books.

Select primary studies
After performing search in selected search venues, potentially relevant candidate studies are obtained that need to be assessed for their actual relevance to the defined RQs. The study selection criteria identify primary studies that provide direct evidence about the research topic under observation. Researchers review specific parts of the articles, such as, title, abstract, keywords for the first stage of screening. To finally decide whether the study is to be included or not, full-text of the candidate papers must be analyzed. To minimize the probability of bias, selection criteria are defined in the review protocol but can be refined during the search process as well. The IC and EC must be appropriate to retrieve all relevant studies from the data sources. Some generic IC such as language, year range of publication, field of interest can already be defined when searching in the electronic libraries.

Assess study quality
To determine the validity of results obtained through the search process, quality assessment is performed. Along with IC and EC, quality assessment criteria are also defined in the review protocol. The quality of individual primary studies is assessed according to a defined quality assessment criteria checklist. Quality scores are used to determine the importance of individual studies when interpreting the findings of an SLR. Based on the quantitative or qualitative nature of the studies, quality criteria checklists are prepared. Assessing the quality of included studies is a challenging task, as there is no standard definition of study quality present in literature. Many of the guidelines and procedures recommended by researchers for assessing study quality, indicate that quality refers to the extent to which the design and execution of an SLR minimizes bias and maximizes validity. The assessment is usually performed manually by the individual reviewers, each study is scored against each criteria. If the opinions of members differ then the methods defined to resolve issues are followed to find unbiased results.

Extract relevant information
After the studies are finally selected, the next step during execution phase of an SLR is to extract relevant information from them. The data extraction forms are designed to accurately retrieve data required to answer defined RQs. To reduce the opportunity for bias, data extraction forms and procedures are usually specified in the review protocol. The design of data extraction forms must be appropriate for the study type included in an SLR. Information to be recorded in the forms include standard publication details, such as, name, date, year, author, publisher, space for additional notes, along with other data to answer the RQ. For qualitative studies, extracted data is usually in textual form whereas for quantitative study type, numerical data is extracted. The researchers fill the forms for each primary study either individually or in groups. The extracted data must be then compared and disagreements, if any, should be resolved either by consensus among researchers or arbitration by an external independent researcher.

Synthesize information
The last step of conducting phase of an SLR, involves applying methods to combine and summarize results of the primary studies. The data synthesis procedure is defined in the review protocol. Similar to the previous steps, the data synthesis approach is also based on the quantitative or qualitative nature of the primary studies. Using statistical techniques to obtain a quantitative synthesis is referred to as *meta-analysis*. Quantitative synthesis requires extracted data to be stored in a way that allows data from different studies to be compared, such as, tables. On the other hand, qualitative synthesis involves integrating results and conclusions explained in natural language. For SE domain, primary studies are more likely to provide qualitative results hence, qualitative synthesis is usually performed. The researchers might use terms and concepts with different meanings in their study [Schröter et al., 2017b]. Therefore, the synthesis of textual data must be performed carefully by the reviewers to answer the RQs.

2.1.3 Phase 3: Reporting

The final phase of an SLR involves documentation of the results and distribution of knowledge to the intended audience in an appropriate manner. It is important for the researchers to precisely describe their contribution to a specific research area and summarize their findings with valid justifications. As mentioned previously, an SLR can be reported in many different ways, such as, technical report, conference paper, journal paper or thesis chapter. The complete details regarding research process, outcomes of each stage and the overall findings must be included while reporting an SLR. As part of the reporting phase, the steps to be followed include, planning and writing a review report as well as validating the report.

Planning and writing review report
This step involves specifying the intended audience who would benefit from conducted research and the most suitable method for distributing knowledge. The potential audience of an SLR include both, practitioners and researchers. Practitioners may be more interested in the practical implications of the review for their own practice, while researchers find the details of methodology and questions arising for future research more interesting. Therefore, it is important to communicate the results of an SLR effectively, the researchers usually plan the final reports during the preparation of review protocol. The reporting of an SLR includes sufficient details regarding the research method employed, so that other researchers may replicate results by following the same procedure as described. Structure of the report is usually well-defined but there may be some variations according to the study conducted. The most common outline includes the abstract, introduction, background of the topic, research method employed, the execution of review plan, results obtained and the analysis performed. Furthermore, the discussion about results and final conclusion drawn are also an important section of the review report.

Validate review report
After the research process used to conduct an SLR is documented, the report must be

evaluated to ensure validity of the method. The review team members must review the reporting of an SLR, to verify that the RQs are clearly specified, the complete research methodology is correctly reported, appropriate data extraction and synthesis procedures are performed. Moreover, it must be verified that results and conclusions are described clearly to answer the RQ. If an expert opinion was asked regarding the review protocol earlier, then the same expert panel could be requested to evaluate the SLR report. It is also possible that other researchers working in the specific research area or SLR methodology review clarity and consistency of the report.

The SLR phases and their stages we describe above are based on the first adapted guidelines by Kitchenham and Charters [2007] for SE researchers to conduct an SLR. Majority of the systematic reviews reported follow these guidelines for planning, conducting and reporting.

2.2 Text Mining for Article Selection

Knowledge Discovery in Databases (KDD) refers to the overall process of discovering high-level, potentially useful knowledge from low-level data [Keim, 2002]. The process involves using database along with any required selection, pre-processing, sub-sampling, transformation, applying data mining methods to enumerate patterns and interpretation/evaluation of results obtained through the mining algorithms. The complete KDD process is summarized in Figure 2.2.

Data Mining is one of the components of KDD process, for extracting new, important, valid and potentially useful information from data [Witten et al., 2011]. It transforms the extracted information into an understandable structure for further use. The idea is to build computer programs that explore large amounts of data and discover interesting patterns within them [Kroeze et al., 2003]. As shown in Figure 2.2, data mining consists of applying data analysis and discovery algorithms to produce a particular enumeration of patterns over the data.

Text mining is similar to data mining, the major difference is that it focuses on extracting small pieces of information from natural language text instead of structured data, to identify relevant and previously unknown knowledge [Hearst, 1999; Kroeze et al., 2003]. It aims to enable users to collect, interpret, curate, and efficiently discover knowledge required for education or research [Thomas et al., 2011]. Text mining is the process of automatically retrieving useful information from unstructured or semi-structured text through identification and exploration of patterns [Gupta and Lehal, 2009; Mooney and Nahm, 2005]. The input text is structured to derive patterns within the structured data as well as interpretation and evaluation of the output [Witten, 2011]. The process comprises of activities such as, information retrieval to retrieve text relevant to the user's query, lexical analysis to study word distributions, pattern recognition, information extraction to identify and extract fragments of text related to the query, data mining techniques to identify links among the pieces of extracted information and visualization of the information [Thomas et al., 2011].

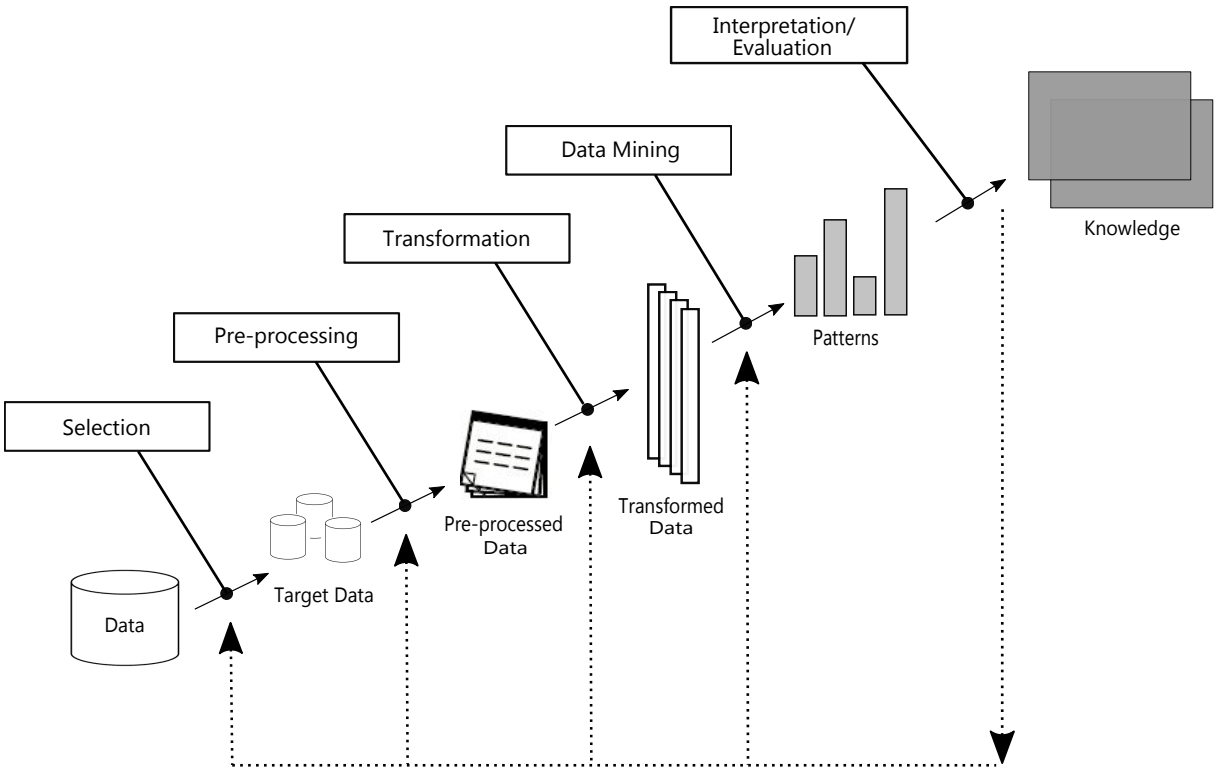

Figure 2.2: Overview of the steps constituting the KDD process based on Fayyad et al. [1996].

Text Mining allows large volumes of diverse information to be merged and managed efficiently [Witten, 2011]. Therefore, significant amount of research regarding text analysis have been conducted by SE researchers to design strategies that support various stages of an SLR, such as, selection of primary studies, quality assessment and data extraction [Felizardo et al., 2012; Kitchenham and Brereton, 2013]. Hence, in this section we will discuss some text mining concepts useful for designing efficient methods to summarize the data about a topic of interest.

2.2.1 Text Mining Methods

For an SLR, large amounts of information must be evaluated to draw relevant conclusions. There are many text mining methods and related tasks that allow efficient analysis of text documents. In this section, we discuss some of the most important methods that support identification of relevant studies in the SLR process. For instance, information extraction, automatic term recognition, document classification/categorization, clustering, information visualization, and text summarization are described below.

Information Extraction

The analysis of unstructured text to automatically extract structured information, such as entities, relationships between entities and attributes describing entities, is referred to as information extraction [Fan et al., 2006]. The process involves identification of

certain small-scale structures or patterns in the text and the relationships between them [Piskorski and Yangarber, 2013]. Once the software infers relationships between the identified patterns, meaningful information is provided to the user. Most of the text mining softwares use information extraction as the basis for many other text mining technologies. This technology is useful for exploring large amounts of text, such as, for an SLR and provide relevant information for further usage [Fan et al., 2006].

Automatic Term Recognition

SLRs are mostly focused on one or more topic areas that can be summarized based on a network of concepts and associations. These concepts can be acquired from the text as technical terms to classify specialized knowledge [Thomas et al., 2011]. Automatic term recognition automatically identifies and extracts keywords within the text and display to the user in a meaningful manner [Ananiadou and Nenadic, 2006]. Keywords are a set of significant words in an article that gives high-level description of its contents to the user [Gupta and Lehal, 2009]. Identification of keywords enables the user to manipulate large volume of text efficiently. This technique is useful for robust processing as it provides the ability to handle unknown words, build lexical and knowledge resources as well as maintain consistency avoiding terminological confusion [Hadjitofallis, 2010]. The automatically generated terms usually correspond to domain concepts which are found in thesauri, controlled vocabularies, and ontologies [Thomas et al., 2011].

There are different methods for automatic term recognition, for example, rule-based approach, dictionary-based, statistical, and machine learning techniques, as discussed by Ananiadou and Nenadic [2006]; Krauthammer and Nenadic [2004]. These methods include the recognition of boundaries of multi-word terms, handling of term variations, such as, acronyms and orthographic variants, to improve the results obtained using these techniques [Thomas et al., 2011].

Document Classification/Categorization

Automatic document classification aims to identify the main theme of a document based on its underlying patterns and distinguishing features, to make them part of a defined group or class [Ananiadou et al., 2009; Yang and Pedersen, 1997]. Using this information, each new document is assigned to the known classes. Categorization counts only words that appear in the text to identify the main topic discussed in the document [Fan et al., 2006]. The classification often relies on a thesaurus for which topics are predefined and relationships are identified based on broad terms, narrower terms, synonyms and related terms. This method can reduce the time required by the user to identify and filter articles relevant to a specific query during an SLR [Thomas et al., 2011].

Clustering

The goal of document clustering is to assign documents to groups based on conceptual similarities [Ananiadou et al., 2009; Fan et al., 2006]. The groups or clusters are generated based on the content and underlying themes shared by all the documents included in that group. A basic clustering algorithm use the words of the documents and their frequencies of occurrence to create a vector of topics for each document. The docu-

ment similarities are calculated based on the generated document vectors [Ananiadou et al., 2009]. Automatic clustering works better with larger document collections, such as, for an SLR, allowing for a more complete view of the domain. Some approaches also provide meaningful labels to the clusters allowing users to efficiently narrow the documents by identifying topics relevant to the search [Thomas et al., 2011].

Information Visualization

The basic idea of data visualization is to present large data sets in some visual form, allowing the user to get an insight into the data, draw conclusions and directly interact with the data [Keim, 2002]. Visual Text Mining (VTM) puts large textual sources in a visual hierarchy or map and provides browsing capabilities to support efficient exploration of documents. For an SLR, visualization tools are used for applying text mining algorithms and methods combined with interactive visualizations to help user select the relevant studies without actually reading their complete text [Malheiros et al., 2007]. For example, a 2D document map shown in Figure 2.3, represents the documents by circles that are placed in a way that similar documents are close together while the ones not similar are placed apart from each other [Felizardo et al., 2011].

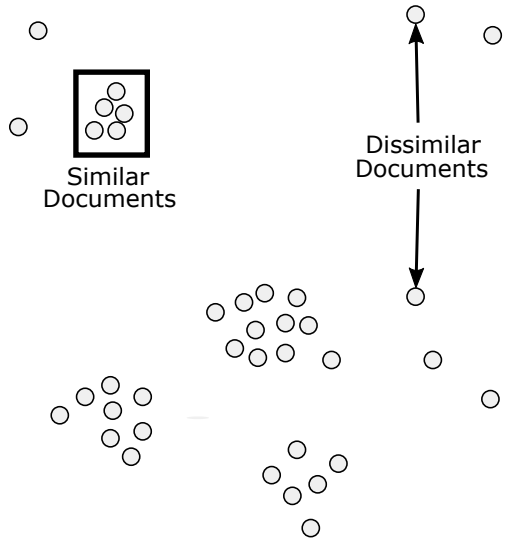

Figure 2.3: 2D document map based on Felizardo et al. [2011].

Text Summarization

The goal of text summarization is to reduce the length and detail of a document while retaining its main points and overall meaning [Fan et al., 2006]. Automatic text summarization is a technique that generates a summary for a document by extracting the most salient information from it [Mani, 2001]. The output of text summarization is intended to be human-readable, allowing the user to determine the relevancy of a document and whether it must be read completely. To develop a software that could automatically analyze semantics and interpret meaning is the main challenge for text mining application developers. One of the strategies commonly used by text summarization tools is

sentence extraction, where important concepts are extracted from the text based on the assumption that a human reader splits each sentence into a set of information fragments to which the sentence refers [Radev and McKeown, 1998]. These information fragments are mutually independent and each of them has an importance score [Thomas et al., 2011]. Summarization tools can also search for specific terms, headings and other markers of subtopics in order to identify the document's key points [Fan et al., 2006]. Many text summarization tools also allow users to choose the percentage of the total text to be extracted as a summary of the document [Witten, 2011].

2.2.2 Text Analysis Process

For applying text mining techniques, the semi-structured/unstructured text must be processed into an acceptable form for efficiently determining the information presented in the document [Gupta and Lehal, 2009]. The collection of documents are pre-processed to a unified format that can be transformed into vector representation Weissi et al. [2010]. Process starts with the collection of studies, each of them is retrieved and checked for its format using a text mining tool. The required information is extracted using text analysis techniques and placed in a management information system. Each of the steps in a text mining process are further explained below according to Weissi et al. [2010]:

Document collection and standardization:
To undergo mining procedures the user must collect the data first. As the documents are selected from different sources, their format can vary. Hence, all the documents must be converted into a standard form for processing. The raw ascii version of the articles are further processed to build numerical vectors, the later steps are discussed further.

Tokenization:
During this step, the unstructured text is examined to identify useful features. Firstly, the character stream is split into words which are called *tokens*. This step is significant to extract higher-level information from the working text. Every single token is a part of a category called "type", so certainly the frequency of number of tokens is much higher than the frequency of types. This process is complicated for a computer program as it does not understand the language structure in a way that human does. Therefore, various conditions and situations are defined for the computer program in order to treat characters as tokens or delimiters. In order to attain optimum results it is better to customize the tokenizer for the working text. This should be noted that the tokenization process is language dependent, therefore the general procedures will remain the same but the details will differ according to the language used.

Lemmatization:
After the process of tokenization, the next step is standardization of tokens that is usually referred as lemmatization or stemming. The goal of stemming is to reduce inflectional and derivational variant forms of a word to a common base form [Goldsmith et al., 2001]. It is an optional but beneficial step for document classification, as it reduces

the number of distinct type in a text corpus and increases the frequency of occurrence of some individual types. For example, an instance "cars" will be reduced to the stem "car" and would be counted as an instance of that type. To perform stemming there are two possibilities and can be selected based on the application.

1. Inflectional Stemming; in English and many other languages, words exist in texts in more than one form. Inflectional stemming is the normalization of these forms on the basis of grammatical variants such as singular/plural and present/past.

2. Stemming to a Root; for some cases of text-processing applications, more intensive stemming is beneficial where the intend is to reach a root form with no inflectional or derivational prefixes and suffixes. For instance, "denormalization" is reduced to the stem "norm". Through such aggressive stemming, the number of types in a text collection is reduced drastically, therefore making distributional statistics more reliable.

Vector Generation (Categorization):
Tokens or words are the main characteristic features for categorizing a document. Without reading the complete content, documents can be described by features that represent the most frequent tokens. The collective set of features is known as *dictionary* that contains words or tokens corresponding to the document collection and forms the basis for constructing a numeric data collection. With the help of this data model, the presence of a word from the dictionary can be checked in a document. This data model can also work along with predictive methods. Frequency information on the word counts improve the predictive performance for some methods. The most frequent but not interesting words, known as "stopwords", are ofter removed from the text before determining the frequency counts. The counts are then modified by the perceived importance of each word by computing weightings for words using the "tf-idf" formulation.

As shown in Equation 2.1, the tf-idf weight assigned to word j is the term frequency (i.e., the word count) modified by a scale factor for the importance of the word. The scale factor is called the *inverse document frequency* and is given in Equation 2.2. It simply checks the number of documents containing word "j" (i.e., df(j)) and reverses the scaling. Thus, when a word appears in many documents, it is considered unimportant and the scale is lowered, perhaps approximately zero. Whereas, if the word is relatively unique and appears in few documents, the scale factor zooms upward because it appears important. As a result, positive scores replace the simple frequency counts which are then transformed into sparse vectors.

$$tf, idf(j) = tf(j) * idf(j), \tag{2.1}$$

$$idf(j) = log(\frac{N}{df(j)}) \tag{2.2}$$

Parsing:
The most sophisticated type of text processing, is producing a full parse of a sentence. Parsing, also known as syntax analysis, is the process of analyzing a string of symbols in natural language, conforming to the rules of formal grammar. The parsing rule is a set of instructions defined by the user to extract relevant data from the text documents. Through this step, the relation and function (e.g. subject, object, etc.) of each word in a sentence with the others is determined. There are different kinds of parsing methods, each associated with linguistic theory of languages.

Feature Generation:
This method is to identify features that can be useful for text mining. Feature generation is accomplished through contextual analysis of the text documents. A feature generator analyzes the documents and maps them onto appropriate ontology concepts, which in turn induce a set of generated features [Gabrilovich and Markovitch, 2005]. In this step, the goal is to reduce the dimensionality and improve accuracy of determining the information present in the text. Through feature generation, it can be determined how well the features depend on the type of object to be classified and also how such features can be obtained.

3. Literature Review

This chapter describes the implementation of an SLR that forms the basis for further work. We conduct the study to identify existing approaches that support the selection activity of an SLR in the field of SE. The goal is to summarize existing approaches to support the SLR process. In particular, we focus on semi-automatic techniques for the selection of primary studies. Through this study, we summarize the existing approaches and identify gaps in the current research. The methodology and execution of the SLR is discussed below. Based on them, we evaluate the results obtained to answer the defined RQs.

3.1 Research Method

We follow the guidelines proposed by Kitchenham and Charters [2007] to conduct an SLR, the steps are summarized in Figure 2.1. To assure transparency and repeatability of the study, we document the complete research process. We define the RQs, search strategy and selection criteria during the planning phase. In the conducting phase, we identify all the available relevant studies and evaluate them against certain selection criteria. We critically analyze the selected studies to find the most relevant and valid primary studies. Finally, we synthesize information extracted from the selected studies to answer the previously defined RQs.

3.1.1 Research Questions

The systematic review process begins with the construction of the RQs that serves as a basis for the study being conducted [Boell and Cecez-Kecmanovic, 2015; Kitchenham, 2004; Kitchenham et al., 2015]. It is the most important step in the planning phase of an SLR, as discussed in Section 2.1.1. The aim of this SLR is to summarize methodologies proposed by SE researchers to support the primary study selection activity in an SLR. Thus, the following RQs are of interest:

RQ1: Which approaches to assess the quality of primary studies have been proposed?
At first, we aim to provide an overview on approaches for semi-automatically assessing the quality of primary studies. Here, we summarize these approaches and describe different aspects, such as, the underlying technique, conducted evaluation, and results. Thus, we support practitioners and researchers that search for an approach to apply or scope a new one.

RQ2: Which underlying techniques are used for this assessment?
For this RQ, we will examine the underlying techniques in more detail. We identify the inputs, processing, and outputs of each identified technique to compare them. Thus, it is possible to identify the techniques that may be more suitable for certain use cases or can be combined with each other.

Overall, we provide a detailed overview on semi-automatic approaches to assess the quality of primary studies in SLRs through this study.

3.1.2 Search Strategy

After formulating the RQs, we define a search strategy to identify all the relevant literature. The strategy includes identifying the keywords for the search and using them to construct search strings. The search process followed for this SLR is an automated keyword search followed with forward and backward snowballing. To increase the chances of obtaining maximum relevant studies, we select primary sources of scientific papers in SE community as the data sources. The search strategy followed for conducting the review is as follows:

Search Terms In order to obtain as many relevant literature as possible, generic and broad search terms are extracted from the review questions. Hence, the main keywords identified in the defined RQs are: “systematic literature review”, “literature quality”, “quality assessment”, “data mining”, “recommender system”, “relevancy categorization”.

Search String To construct the search strings we identify the synonyms, variations and related words for the defined keywords. The search query is a combination of the keywords using the logical operator AND with the synonyms, variations and related terms using the logical operator OR. Using the appropriate boolean expressions the following search string is produced for the review to obtain maximum possible relevant results:

(approach OR support OR method) AND (“systematic literature review” OR “systematic review” OR “systematic literature reviews” OR “systematic reviews” OR SLR) AND (“quality assessment” OR “literature quality”) AND (“data mining” OR “recommender system”) AND (“relevance categorization” OR “relevant study”)

Data Sources After defining the search strings, we select search venues for the SLR. Electronic databases with a large number of SE publications are chosen. Table 3.1 gives the list of data sources we use to identify studies for the review. Furthermore, we carefully study the resulting articles and filter them according to the defined study selection criteria. Finally, we perform forward and backward snowballing to identify studies that might have been overlooked in the initial search.

Database	URL
ACM Digital Library	*http* : *//portal.acm.org*
IEEE Explore	*http* : *//ieeexplore.ieee.org*
ScienceDirect	*http* : *//www.sciencedirect.com*
SpringerLink	*http* : *//www.springerlink.com*

Table 3.1: Data sources for the SLR.

3.1.3 Study Selection Criteria

To identify relevant primary studies to address the review questions, we define the study selection criteria. We use the IC and EC to select papers from the search results of the data sources as explained in Section 2.1.2. The IC and EC defined for the literature review we conduct are listed below:

Inclusion criteria

- The study must address an approach focusing on quality assessment or relevancy categorization in the SLR process.
  Different automated approaches have been proposed by researchers to support the stages of an SLR. This review will focus on methods that assist selection and quality assessment of primary studies.

- Article elaborates the methodology used to support the approach.
  Sufficient information for the reader to clearly understand the steps followed for achieving the results is provided in the study.

- Study is conducted between 2007 to 2016.
  The guidelines to lead an SLR in SE were proposed by Kitchenham and Charters [2007]. This review targets articles that are published after these guidelines and until the time we conduct the review.

- Paper reviewed and officially published in a journal/workshop/conference.
  To ensure good quality results, the proposed approach must be reviewed prior to the publication.

- Articles related to SE domain.
  SLRs have been widely used as a research method in other domains such as medicine and social sciences. This criteria is to limit our focus on studies conducted in the CS domain.

Exclusion criteria

- Papers not written in English.
  A standard language is selected as a means of representing information.
- Abstracts and Power Point presentations.
  To obtain detailed description of the proposed approach, complete text file of the study is preferred that provides the required amount of data.
- Technical reports and Bachelor/Master/PhD theses.
  The articles published by experts in the field of CS are more likely to be detailed and reviewed.
- Articles with unknown publication type or publisher.
  To obtain better quality results, articles with incomplete information regarding authors or publishers are excluded.

3.1.4 Quality Assessment

We assess the selected primary studies based on defined quality criteria to increase reliability of the results. We use the guidelines suggested by Kitchenham et al. [2015] and Kitchenham and Charters [2007] for performing rigorous quality assessment of the selected papers. To assess the selected articles manually, the quality criteria checklist contains the following questions:

$Q1$ Is there a clear statement defining the objective and aim of the research?
The goal of research carried out in the study must be clear, if there are unclear descriptions the paper is given a score 0=No, while if there is only partial description of the objectives the score given is 0.5=Partial, otherwise the score is 1=Yes.

$Q2$ Is there an adequate description justifying the choice of research area?
Studies with an explanation to support selection of the topic for research are assigned a score 1=Yes, if not 0=No.

$Q3$ Is the research method appropriate to address the aims of the research?
The authors must explain clearly and completely the proposed strategy. This includes elaboration of all the intermediate steps and their purpose. If relevant, clear and adequate descriptions are provided the paper is given a score 1=Yes, while for partial and unclear descriptions the scores given are 0.5=Partial and 0=No respectively.

*Q*4 How feasible is the method followed to achieve results?
The results of the feasibility study conducted represent how promising the proposed approach would be. The limitations to the approach, if any, must be identified by the researchers. If the proposed approach is useful for the purpose the score 1=Yes is assigned, while if the approach is partially useful then 0.5=Partial is assigned.

*Q*5 Is there clear and coherent reporting of the findings?
The paper must provide valid results obtained by applying the proposed strategy. If a relevant feasibility study is described by the researchers to evaluate the proposed approach the score assigned is 1=Yes. If there is no study performed to evaluate the proposed strategy the assigned score is 0=No.

*Q*6 Is the study of value for research and practice?
Studies summarizing whether the findings of the study can benefit SE researchers in practice are assigned a score 1=Yes, otherwise 0=No.

The quality assessment criteria we list above cover four main areas of empirical research as explained by Dybå and Dingsøyr [2008b]:

- *Reporting:* criteria Q1 and Q2 relate to the quality of reporting an empirical study based on the aims and description,
- *Rigor:* criteria Q3 and Q4 address how appropriate and thorough the applied research design is for achieving valid results,
- *Credibility:* criteria Q5 focus on assessing the reliability of findings concluded through the study,
- *Relevance:* criteria Q6 concerns the relevancy of the study for the research community.

For assessing quality, we carefully analyze the selected primary studies. We assign each primary study a score according to the above mentioned quality criteria. For each of the quality criteria the scores are given on the scale of 0 to 1 with maximum score representing the best quality paper. We discuss the scores assigned to each of the primary studies for this review in Section 3.2.3.

3.1.5 Data Extraction and Data Synthesis

After the selection and quality assessment of primary studies, we extract relevant information from them to answer the defined RQs. Through this step, we aim to acquire data from individual primary studies and synthesize in a manner appropriate for our study. We design the data extraction forms to accurately record the information obtained from the primary studies. In addition to including the RQs and study selection criteria, we record the following standard information for each paper, as suggested by Kitchenham and Charters [2007]:

- Primary study ID.
- Author(s).
- Title.
- Publisher.
- Year of publication.
- Publication details (volume and issue).
- Page numbers.

Furthermore, we extract the primary study specific data for each paper to answer the review questions. To keep the extracted information consistent, we collect the following information for each primary study:

- Summary of aims of the study.
- Name and description of the method proposed.
- Explanation of methodology used to achieve the defined goals.
- Summary of findings.
- Results of evaluation, if performed, for the proposed strategy.

We study each of the primary articles carefully to identify the basic concepts, methodology followed and conclusions drawn regarding the proposed approach. We synthesize the data extracted from the primary studies to answer both of the defined RQs, discussed later in Section 3.3.

3.2 Conducting the Review

After the planning phase, we conduct the review following the defined methodology. We obtain the review plan to perform the SLR as a result of first phase. The conducting phase involves identification of relevant research, selection of primary studies and quality assessment. In the next part of this section we discuss each of the stages involved to execute the SLR:

3.2.1 Identifying Relevant Research

The conducting phase begins with identifying relevant studies from the selected data sources. The search results retrieve a total of 3,196 potentially relevant articles for the search query. The Table 3.2 shows the number of articles we initially obtain from each of the selected data sources. The results obtained through the initial search can be found in Section A.1.

Data Source	Search Result
ACM Digital Library	2076 articles
IEEE Xplore	31 articles
Science Direct	474 articles
SpringerLink	615 articles

Table 3.2: Search results of the selected data sources.

3.2.2 Selection of Primary Studies

We screen the search results obtained from the selected data sources based on title, abstract and keywords for each paper. For a secondary review, the selection is related to primary studies hence, any secondary or tertiary study are excluded during the initial screening [Octaviano et al., 2015]. The primary studies we select to answer the defined RQs are listed in Table 3.3.

Year	Title	Library	Cited as
2007	A Visual Text Mining approach for Systematic Reviews	IEEE Xplore	[Malheiros et al., 2007]
2011	Linked Data approach for Selection Process Automation in Systematic Reviews	IEEE Xplore	[Tomassetti et al., 2011]
2011	Using Visual Text Mining to Support the Study Selection activity in Systematic Literature Reviews	IEEE Xplore	[Felizardo et al., 2011]
2015	Applying Information Retrieval Techniques to Detect Duplicates and to Rank References in the Preliminary Phases of Systematic Literature Reviews	SemanticScholar	[Abilio et al., 2015]
2015	Semi-automatic Selection of Primary studies in Systematic Literature Reviews: is it reasonable?	SpringerLink	[Octaviano et al., 2015]
2016	Improvements in the StArt Tool to better Support the Systematic Review Process	ACM	[Fabbri et al., 2016]
2016	PaperQuest: A Visualization Tool to Support Literature Review	ACM	[Ponsard et al., 2016]

Table 3.3: Selected primary studies.

To obtain relevant primary studies we prepare an outline of the required steps, given in Figure 3.1. We initiate the search and selection process by screening the title, abstract and keywords of the 3,196 potentially relevant studies. After obtaining a set of 11 selected candidate papers, we perform forward and backward snowballing. As a result, we retrieve 18 studies and analyze their full-text based on the study selection criteria. Finally, we select 7 primary studies to answer the defined RQs.

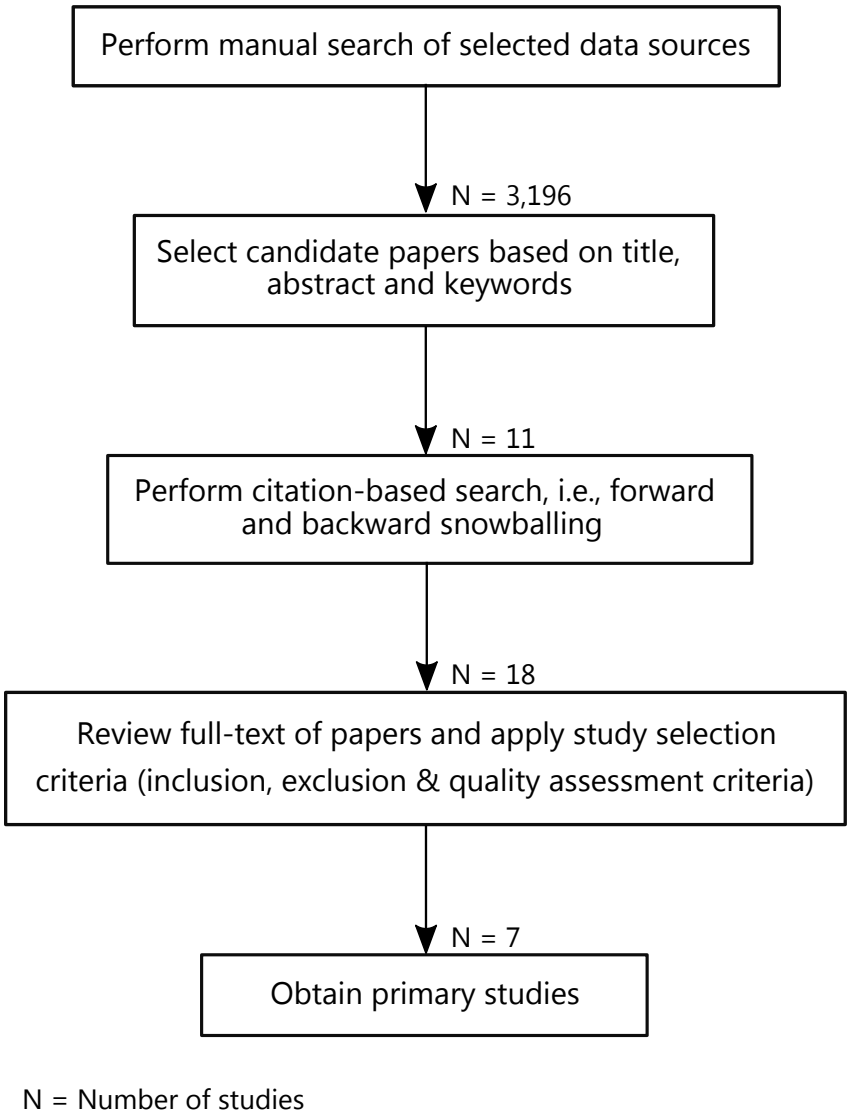

Figure 3.1: Procedure followed for identifying primary studies.

3.2.3 Quality Assessment

The last step of the conducting phase is the quality assessment of selected primary studies. Using the checklist defined in Section 3.1.4 we assign scores to each primary study. We assign scores separately for every quality criteria and then combine them to form the final score. The scores given to the selected relevant studies are shown in Table 3.4.

We score each of the quality assessment criteria according to the procedure followed in Kitchenham et al. [2009]. The scoring procedure is 1=Yes, 0.5=Partial, 0=No. We sum the individual score for the studies and assign on the scale of 6.

Primary Study	Quality Assessment						Score (/6)
	QC 1	QC 2	QC 3	QC 4	QC 5	QC 6	
[Malheiros et al., 2007]	1	1	1	0.5	1	1	5.5
[Felizardo et al., 2011]	1	1	1	0.5	1	1	5.5
[Abilio et al., 2015]	1	1	1	0.5	1	1	5.5
[Tomassetti et al., 2011]	1	1	0.5	0.5	1	1	5.0
[Octaviano et al., 2015]	1	1	0.5	0.5	1	1	5.0
[Fabbri et al., 2016]	1	1	1	0.5	0.5	1	5.0
[Ponsard et al., 2016]	1	1	1	0.5	0.5	1	5.0

Table 3.4: Quality assessment of primary studies.

Observing the assigned scores in Table 3.4, we conclude that all the strategies proposed by SE researchers to support primary study selection activity only partially satisfy Q4. The feasibility study performed to evaluate the proposed strategies indicate some limitations, discussed by the researchers as threat to validity. For example, Malheiros et al. [2007] and Felizardo et al. [2011] perform the case study with a limited article sample, hence their strategy could not be validated for real SLRs. Also, the participants involved in the experiment are limited and results are influenced by their individual level of expertise.

The strategy proposed by Tomassetti et al. [2011] involves construction of an initial set of relevant papers based on researcher's previous knowledge regarding the topic, hence causing subjective bias. Moreover, this initial set of papers represents only a small part of literature available regarding the topic of interest. As a result, automatic classification could potentially discard resources not part of the initial set. Furthermore, a construct threat in the Linked Data enrichment step and a conclusion threat in the Naïve Bayes classification are also possible drawbacks of the proposed approach.

Similarly, the strategies recommended by Octaviano et al. [2015] and Abilio et al. [2015] analyze part of content for scoring the articles; title, abstract and keywords. The strategies are not evaluated to analyze full-text of papers to confirm their reliability. The Score Citation Automatic Selection (SCAS) strategy proposed by Octaviano et al. [2015], implemented in the StArt tool discussed by Fabbri et al. [2016] considers an article with a high score, assigned based on comparison of search string terms with the text, receiving at least one citation as an included primary study. There could be a possibility that a recent relevant paper not cited, is excluded from the study according to the SCAS recommendation. Furthermore, the search queries need to be constructed carefully to obtain reliable results for the proposed approaches.

Ponsard et al. [2016] present a visualization tool to support efficient reading decisions, not formally evaluated. The algorithm employed considers citation links of the papers as the only criteria for determining relevancy.

3.3 Results

In the following section of this chapter we discuss results obtained from the SLR conducted. We answer each RQ based on our findings from the selected primary studies given in Table 3.3.

RQ 1) Which approaches to assess the quality of primary studies have been proposed?
As a result of the SLR we obtain 7 primary studies, suggesting approaches to support the selection of primary studies. Most of the approaches are based on *VTM technique*, using supporting tools, such as, Projection Explorer (PEx) and ReVis for the implementation.

Furthermore, one of the approach is based on *Linked Data and Text Mining technique*. It uses Naïve Bayes classifier to filter potentially relevant articles that can be examined by researchers for the final selection.

Information Retrieval (IR) technique to rank references for an SLR is also proposed by researchers. The ranking is based on relevance between the defined search query and specific parts of text.

Finally, to determine the relevancy of a study their *Citation Links* are suggested as a connectedness measure. The citation relationships of relevant papers are sorted to support efficient reading decisions by the reviewer.

We summarize the approaches, tools used, study type, context and main results in Table 3.5.

RQ 2) Which underlying techniques are used for this assessment?
The different approaches identified in the previous RQ to support selection of primary studies in an SLR are elaborated to answer the second RQ. We describe the techniques based on their usage by the proposed strategies.

Visual Text Mining
Text Mining is a commonly used practice to extract patterns and non-trivial knowledge from unstructured or text documents written in natural language. Interactive visualizations combined with text mining algorithms and methods, support efficient exploration of a large set of documents.
Input: Full text of articles obtained through the search query and converted into raw ascii text format. Additionally, researchers define the exploration strategies to identify relevant documents from the visual representations.

Processing: The VTM tool provides methods to calculate document similarities, following are most likely to be used:

Study	Name	Study Type	Approach	Tools used	Study context and Main results
[Malheiros et al., 2007]	VTM-Based Systematic Review	Exp.	VTM	PEx	Three researchers study 100 articles, participants B and C used VTM 2D document maps, while A did not. Using an oracle of 40 papers: A found 8.67 articles/h, B and C found 24.49 and 23.53 articles/h, with a precision of 82.8% (A), 81.28% (B) and 92% (C).
[Tomassetti et al., 2011]	N/A	CS	Linked Data; Text Mining	DBpedia; OpenCalais Web; Naïve Bayes	Researchers report a process that supports the second phase of data selection based on keywords and a Naïve Bayes classification process. The process was trailed on part of a large cost estimation SLR and amount of papers needing manual review reduced by 20%.
[Felizardo et al., 2011]	SLR-VTM	Exp.	VTM	ReVis	Four PhD students: two participants use VTM diagrams (document map, edge bundle, citation network) had better performance and more reliable outcomes selecting primary studies compared to the two who read the abstracts manually.
[Abilio et al., 2015]	N/A	Exp.	Information Retrieval	Vector Model	Two strategies to rank references are implemented on two real datasets; one of the strategy presented 50% of precision and 80% of recall.
[Octaviano et al., 2015]	SCAS	CS	VTM	StArt;ReVis	Three existing manually conducted SLRs are selected for implementing SCAS strategy to automate initial selection of primary studies. Result: average effort reduction was 58.2% and percentage error was 12.98%.
[Fabbri et al., 2016]	StArt	Features	SCAS	StArt	Researchers implement the SCAS strategy in StArt to semi-automate the selection of primary studies and improve the support provided by the tool for conducting an SLR.
[Ponsard et al., 2016]	PaperQuest	Tool	Citation Links	PaperQuest; Google Scholar	A visualization tool is described that supports efficient reading decisions, by only displaying information useful at a given step of the review. The suggested algorithm will find and sort papers that are likely to be relevant to users, based on papers they have already expressed interest in and their citations.

Exp.= Experiment, CS = Case Study, N/A = Not Available.

Table 3.5: Summary of the proposed approaches supporting primary study selection.

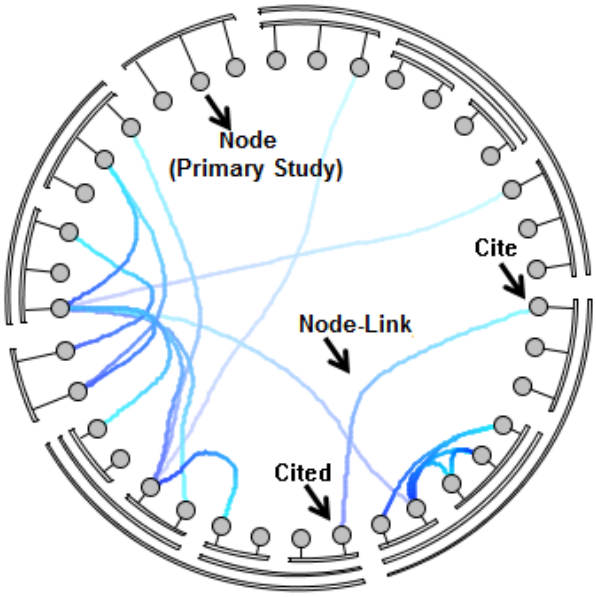

Figure 3.2: Edge Bundle [Felizardo et al., 2011].

i. A vector space model of ascii documents is built from the selected terms. The document x term matrix is filled with 'term frequency, inverse document frequency' measure. The similarity between two documents is then computed as cosine distance between their vectors.

ii. The document x document distance matrix is produced for all raw ascii files. This method is based on Normalized Compression Distance, which is the approximation of conditional Kolmogorov complexity.

Output (a): The VTM tool presents visual representations, such as, 2D document map discussed in Section 2.2.1 (Figure 2.3), edge bundle as in Figure 3.2 and citation network shown in Figure 3.3 to help the user effectively explore a collection of documents without actually reading all of them.

Output (b): Using the combination of two VTM tools to classify primary studies based on relevance of content and citation relationships. The categorization is performed into three categories and four quadrants as shown in Figure 3.4.

Linked Data approach and Text Mining

To assess the similarity of documents and categorize them, automatic text classification technique can be applied. It involves assigning a text document to a category automatically using a machine learning technique. The classification is usually performed based on significant words or features extracted from the text.
Input: An initial set of relevant documents is required as an input along with a set of candidate papers obtained after applying the search query.

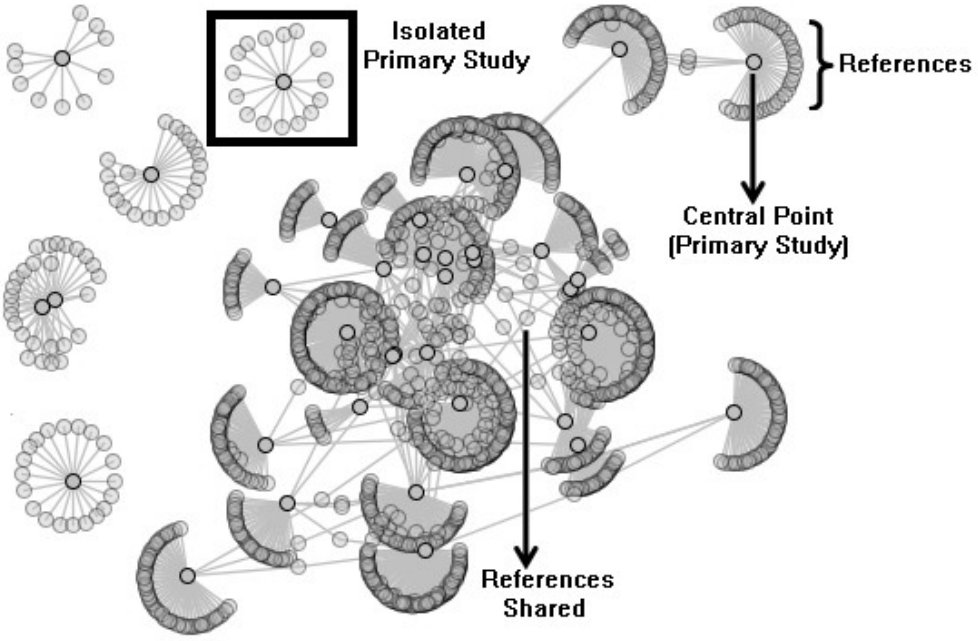

Figure 3.3: Citation Network [Felizardo et al., 2011].

Processing: The pre-processing steps applied to the text documents include removal of stop words and stemming. The processing then involves the following steps:

1. Using the set of documents, feature extraction is applied to identify important words or features in the text document. Features or attributes usually mean significant words, multi-words or frequently occurring phrases indicative of the text category. Mostly, a bag of words model based on specific parts of content is built to make comparisons between the initial set of relevant papers and the candidate papers.

2. After feature selection, the text document is represented as a document vector, and an appropriate machine learning algorithm is used to train the text classifier. Naïve Bayes, Decision Tree, Neural Network, Support Vector Machine and Hybrid approach are a few text classifiers that can be used.

3. Linked Data approach is used for the enrichment of papers, which are then compared with the defined model. The initial set of included papers used to build the model changes after some papers are added to this set as a result of the manual review hence, the model must be re-build.

Output: Text documents classified in the set of relevant papers are then read manually for the final selection as included or excluded.

IR techniques

IR deals with the representation, storage, organization and access to information, such

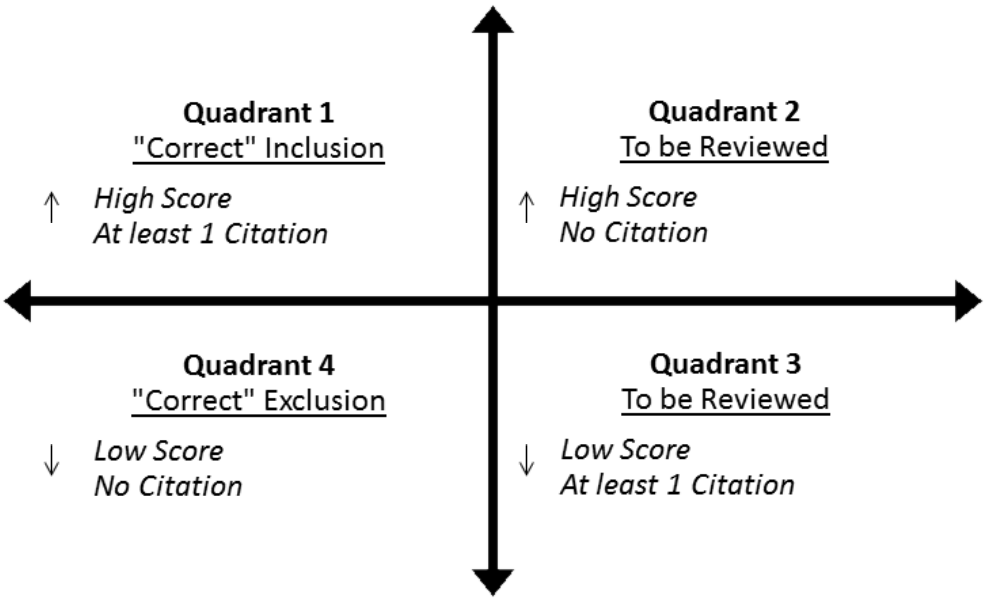

Figure 3.4: Classification of primary studies in SCAS based on Octaviano et al. [2015].

as documents, web pages, on-line catalogs, structured and semi-structured records and multimedia objects. The Vector Model is a classic algebraic model of IR that represents documents and queries as vectors in t-dimensional space (where, t is the number of distinct terms). Aiming to minimize the effort required to manually perform selection of primary studies in an SLR, IR techniques are employed.

Input: A set of references to be analyzed and a search query is given as input. The search query can be treated in two ways:

i. The partial contribution of each query term is considered, where all terms are treated in the same way regardless whether they are synonyms or not. The duplicated terms and connectors, AND and OR, are also ignored.

ii. To simulate boolean expression of search string, the terms are organized in groups that aggregate a term and its synonym using OR connector. These groups are then connected by the AND connector.

Processing: After the pre-processing of the text; removal of stop words, punctuation, symbols and conversion of letters to lowercase, the processing of the input involves the following steps:

1. Distinct token from the articles are inserted into an inverted index structure composed of key-value pairs. Each pair consists of the key as a distinct token and the value as a list of occurrences containing, for each document in which the token occurs, its frequency of occurrence and identification of document.

2. The occurrence of each token of query in the inverted index is verified and weight of each token of query/group in the document is calculated. The similarity of the document with defined query is then determined.

Output: Similarity calculated for each of the corresponding documents are inserted in a list that is sorted in descending order and returned to user.

Citation Links (Information Visualization)
The relevance of an article can be determined by its citation links to articles that the user found interesting. One of the fundamental fact of scientific literature is that authors build their research upon previous works. Hence, by interpreting these citations as links in network, the relatedness between studies can be determined.

Input: A few seed papers that user found relevant to the topic of interest and their citations are used to discover more relevant papers.

Processing: The space of previously published papers regarding a topic is divided into the following subspaces;

a. Core: papers already read, upon which the user builds their understanding.

b. Fringe: papers a researcher has access to because they are referred or cited by some papers from the Core.

c. To Read List: articles the user found interesting from the Fringe but not read yet.

d. Unknown: all the other papers that are not related to the ones found interesting.

A connectedness measure is computed for each paper in the Core, Fringe and To Read List as weighted sum of all links between the paper under observation and all the other papers found interesting. This connectedness measure and citation count from an external library are combined, then normalized to determine the relevance score of an article.

Output: Results obtained through the processing are displayed as a list with a top-down reading order according to their relevance score.

3.3.1 Threats to Validity

In this section, we discuss factors that can effect reliability of the study we conduct. The limitations on the validity of an SLR are expressed as threats to validity [Kitchenham et al., 2015]. Factors that may effect validity of the study may arise in the design, conduct, analysis and reporting of an SLR. The following are four major forms of threats to validity that may effect results of the study we perform:

Construct Validity
The construct validity is concerned with how well results of the study are linked to the

concepts and theory behind it. We formulate the RQs to focus on primary studies that address approaches supporting conducting phase of an SLR. Hence, we conduct this study in general following the review protocol we define earlier to obtain the results.

Internal Validity
The internal validity identifies factors during the execution, that might cause some degree of bias in overall process. The study we perform targets literature available in selected libraries, published in the field of CS during a specific time range. Hence, there can be a certain degree of limitation to internal validity of the SLR we perform, which could not be completely eliminated.

Conclusion Validity
This validity focuses on factors that can effect analysis of the selected primary studies. We analyze each study carefully to draw meaningful and valid conclusions. Results obtained through the search string are reviewed individually by more than one reviewer to minimize errors.

External Validity
The external validity is concerned with how well conclusions from the study may be generalized. The study we perform describes the current state of research. Hence, results we obtain are valid to a limited extent, especially as properties of search engines change rapidly over time and cannot be avoided.

Overall, there are some limitations to the validity of results obtained through the SLR we conduct. However, these are only limited and partly intended for this investigation.

3.4 Summary

This chapter describes an SLR we conduct to analyze the existing approaches supporting the selection of primary studies and quality assessment. The study we perform provides the current state of research to support the SLR process. In this study we discuss 7 primary studies that propose different methods to semi-automate the primary study selection activity during an SLR. We identify a variation in the approaches but overall text mining techniques prove to be useful for accelerating the SLR process. However, we conclude that reliability of the approaches must be further investigated. Most of the approaches perform content analysis based on only specific parts of the text hence, researchers must aim for strategies to analyze the full text of papers for improved ranking. The quality of primary studies define the quality of secondary study being conducted and to achieve reliable results of an SLR, quality assessment of included papers is an important step. The approaches we discuss to answer the RQs focus on selection of primary studies but assessing them for validity lacks tool support. Therefore, in the next chapter we aim to present a methodology that supports quality assessment together with selection of primary studies.

4. Methodology

As we discuss in the previous chapter, existing approaches to semi-automate the conducting phase of an SLR focus mainly on analyzing the documents based on content and citation relationships. In this chapter, we propose an approach to support the selection of primary studies based on quality assessment along with relevancy of the content. We begin with an explanation of requirements for the approach and further on explain the methodology. Furthermore, we discuss some limitations of our approach and the possible solutions to overcome them.

4.1 Prerequisite for the Approach

To support the reviewers efficiently conduct an SLR, we propose an approach to semi-automate the selection and quality assessment of primary studies. For our approach, the first and third phase of an SLR remain exactly the same as described by Kitchenham and Charters [2007] and explained earlier in Section 2.1. We modify the second phase of the process and the most important requirement for our approach is the planning stage. In Figure 4.1 we summarize the prerequisite stages and highlight the steps that our approach aims to support during the second phase of an SLR.

4.1.1 Phase 1: Plan the review

Following the steps described in Section 2.1.1, a review protocol is prepared in the planning phase. The reviewers specify the need to perform a review in the specific research area during this phase. The RQs, search strategy and study selection criteria must be defined before conducting an SLR. Constructing clear and meaningful RQs is one of the most essential requirement for identifying relevant primary studies. Also, to ensure that an SLR focuses on a specific issue, careful formulation of RQs is necessary [Dybår et al., 2007].

Based on the defined RQs, keywords are determined to generate a search query. The extracted search terms are the basis for identifying relevant primary studies. All the

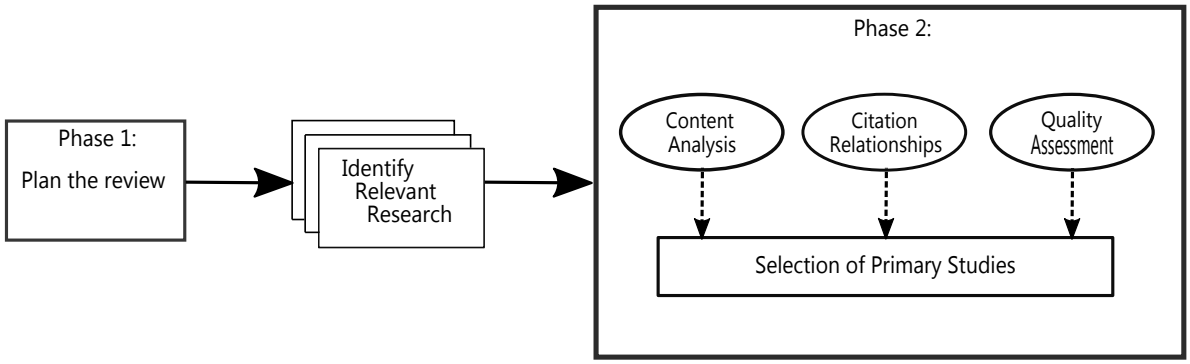

Figure 4.1: Summary of steps based on the approach.

related terms and synonyms of the keywords must be specified to determine all the potentially relevant studies. The search query is structured using the boolean operators, AND and OR to obtain results from the selected data resources.

During the planning phase, the study selection and quality assessment criteria must also be defined. Some of the general criteria such as language, year range, and field of interest can be defined during the search in selected data resources. The other defined IC and quality assessment criteria are applied later in the conducting phase to select the most relevant primary studies. For quality scoring of the articles, we define certain criteria for our approach which is discussed later in this section.

4.1.2 Identify Relevant Research

After planning phase, the review protocol needs to be followed for performing search in the selected data resources. All the available evidence regarding the topic of interest is obtained by applying the search query to the digital libraries, such as, ScienceDirect, ACM, and IEEE Xplore. For our approach, it is required to export the citations and abstract of search results. We export the *BibTeX file format* to obtain bibliographic meta data and abstract of each article.

The most commonly used bibliographic modeling approaches include, BibTeX, EndNote and RIS file formats. Many popular digital libraries and scholarly search engines support exporting citations in these format. BibTeX uses a style-independent text based file format for the bibliographic entries, such as articles, in-proceedings, books, and theses [MacGillivray and Jim Pitman]. The bibliographic information is represented in JavaScript Object Notation (JSON) form that is more compact and can be easily loaded in JavaScript. This data can be shared between web applications to efficiently make better use of the data. The file stores entries in a way similar to a database, consisting of records and fields. Each BibTeX record holds the bibliographic information for a single bibliographic entry [Wenneker, 2010]. Figure 4.2 shows how a bibliographic record would be entered into a BibTeX file. The records begin with an "@" symbol,

followed by the entry type and, in braces, a comma-separated list of entries of the form "fieldname = value", where the 'fieldnames' are components of the bibliography entry, such as, author, title, etc.

```
@article{Eadie2012390,
title = "Recommendations for research design and reporting in computer-assisted diagnosis to facilitate
meta-analysis ",
journal = "Journal of Biomedical Informatics ",
volume = "45",
pages = "390 - 397",
year = "2012",
doi = "https://doi.org/10.1016/j.jbi.2011.07.009",
url = "http://www.sciencedirect.com/science/article/pii/S1532046411001274",
author = "Leila H. Eadie and Paul Taylor and Adam P. Gibson",
keywords = "Computer-assisted diagnosis",
keywords = "Radiological imaging",
abstract = "Computer-assisted diagnosis (CAD) describes a diverse, heterogeneous range of applications
rather than a single entity. The aims and functions of CAD systems vary considerably and comparing
studies and systems is challenging due to methodological and design differences. In addition, poor study
quality and reporting can reduce the value of some publications."
}
```

Figure 4.2: An example of a BibTeX entry.

4.2 Approach for Selection and Quality Assessment

We propose an approach to support the selection of relevant, most promising primary studies for an SLR. As we discuss earlier in Section 2.1.2, manual selection of relevant studies from a large collection of documents requires a lot of time and effort. Therefore, the focus of our approach is to allow reviewer to efficiently select relevant primary studies. We alter the conducting phase of an SLR, an overview of our methodology and details follow in the next subsections. Figure 4.3 summarizes the steps of our approach during conducting phase.

For the approach, user provides keywords for the SLR and search results initially obtained from the data resources. To analyze the content relevancy, title, abstract and keywords of each article is used to determine the frequency count of terms. Furthermore, by exploring citation relationships of the papers, we aim to identify any relevant study that might have been overlooked during the initial screening of articles. The final step during processing is assigning scores to each selected article based on quality assessment criteria. As a result, the list of selected articles with their assigned scores are displayed to the reviewer. The approach we propose is presented in context of the guidelines recommended by Kitchenham and Charters [2007] for SE researchers, however, it can also be applied isolated or together with an SLR in different scientific disciplines to obtain most relevant promising studies.

Input Keyword and Search Results

The web-tool we design to implement our approach requires user to initialize the process by providing keywords. These search terms are used for content analysis of the articles. Each keyword and the related terms are defined individually to count their occurrences in the title, abstract and keywords of each study.

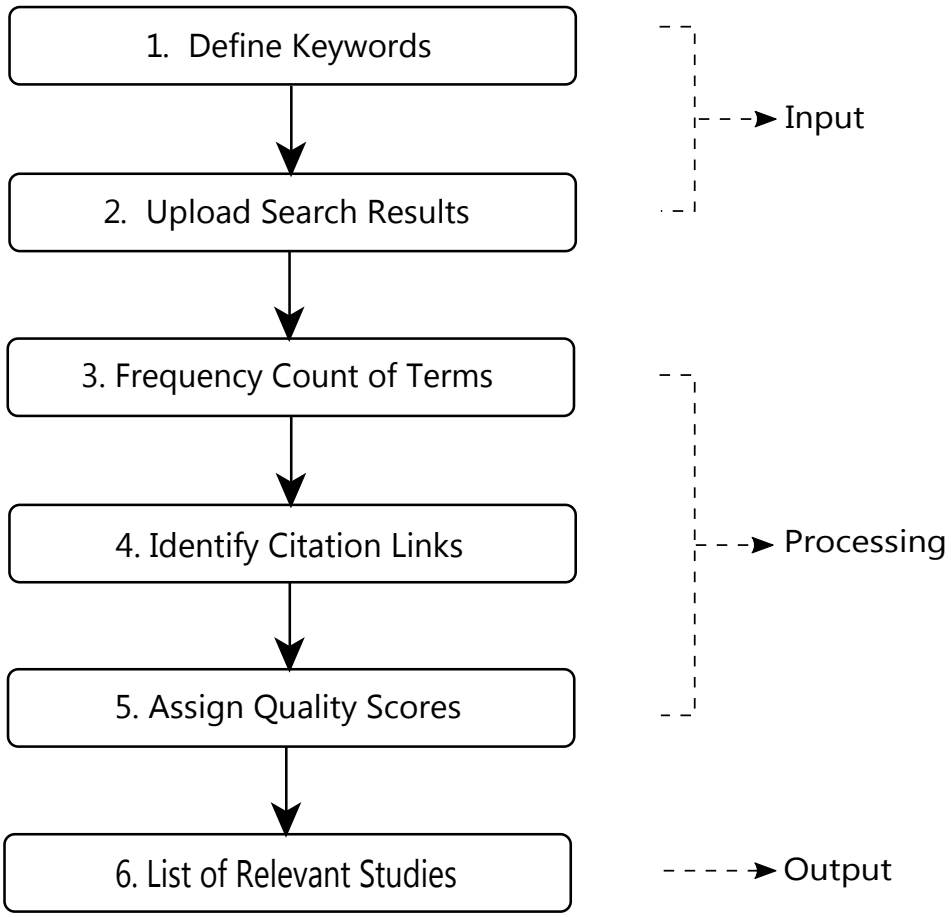

Figure 4.3: Overview of the approach for selection of primary studies.

Moreover, initial search results obtained from the data sources and stored as BibTeX file format need to be uploaded by the user. Bibliographic meta data and abstracts of documents are represented in a form shown in Figure 4.2. The JSON format is text only, that can easily be sent to and from a server using any programming language [W3Schools]. To identify the most relevant primary studies, a text file containing the bibliographic information is loaded in the tool.

4.2.1 Content Analysis

The first phase of our approach involves identification of studies based on their titles, abstracts and keywords. We use *BibTeX parsing* to organize the field names and their values for the entries in a structured format. Each field name is comma separated and their values are present in quotes. The entry type and field names are considered as columns while the values are stored as rows. Once the bibliographic information is organized in a database, we extract the most occurring words and frequency count of defined keywords. Prior to the text analysis, irrelevant terms are removed to reflect meaningful terms present in each document. Thus, the pre-processing involves removal of stop words, such as, conjugations, articles and prepositions.

Each article is evaluated to determine frequency count of the defined keywords and top ten most frequently occurring terms in the text. Analysis of the textual information we propose is outlined in Figure 4.4. The two matrices are displayed to the user to give a good insight about each document. Screening the titles and frequency counts would enable the reviewer to make effective decision regarding their relevancy with the topic under observation.

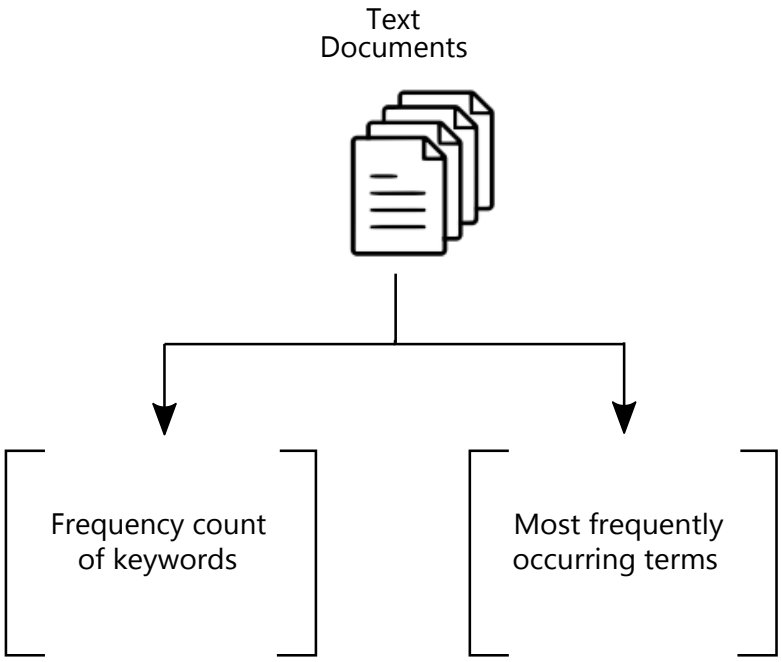

Figure 4.4: Analysis of textual information for content analysis.

4.2.2 Citation Relationships

After performing the content-based selection of articles, the next step of our approach is to determine their related articles as demonstrated in Figure 4.5. Through this step, we aim to identify relevant articles that might have been overlooked during the initial selection. The referred and cited-by articles can be recognized using "doi" and "url" of the selected ones. Using JavaScript, we aimed to access the related articles automatically but due to restrictions imposed by the digital libraries, this remained impossible. Thus, the approach requires reviewer to manually perform backward and forward snowballing. The meta data and abstracts of articles identified through citation relationships need to be uploaded by the user as previously done for the search results. This bibliographic information is stored in a structured form, similar to the first phase, to be used later on for assigning quality scores of the articles.

4.2.3 Quality Scoring

The final stage in processing involves assigning scores to each article based on a set of defined quality criteria. The meta data and abstract of selected articles are analyzed

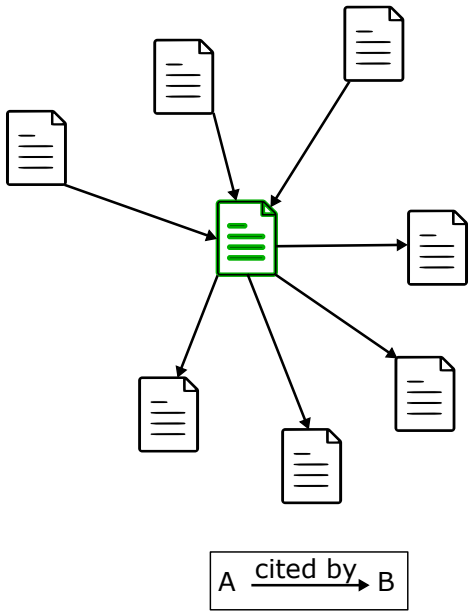

Figure 4.5: Identify related articles, adapted from Lausberger [2017].

to determine the individual scores for each criteria. These are then combined to assign a final score of each study. Some of the quality scoring criteria that we define for our approach are assigned values after *normalization* of the actual value to obtain results between 0 and 1. The equation we use to normalize the values is given in Equation 4.1. To obtain the normalized value, we divide the original value of the score with the maximum score assigned.

$$value_{norm}(j) = \frac{value_{original}(j)}{value_{maximum}} \tag{4.1}$$

For our approach we define the following quality scoring criteria:

QC1. How relevant is the information present in a study to the topic under observation? Based on titles, abstracts and keywords of studies, frequency counts of the user defined terms are determined to analyze the relevancy of the content. Articles with frequent occurrences of keywords are considered relevant to the topic area of interest. The retrieved frequency counts of each study are normalized and studies with less relevancy are assigned scores closer to 0 while relevant studies have higher scores.

QC2. How credible are the findings of an author performing research in a specific area? To determine the quality of a study, it is important to perceive the level of expertise of the authors. Higher contribution of a researcher in a specific research area indicates an expert level. Thus, authors publishing more papers based on the topic of interest are assigned higher scores as their findings are more likely to be

valid. For our approach, the author count is normalized to assign scores between 0 and 1.

QC3. Is the study peer-reviewed and officially published?
Studies that are officially published are reviewed prior to their publication. Hence, their findings are most likely to be accurate and reliable. Articles published in sources, such as, journal, conference or workshops are assigned 1=Yes, while other research articles not officially published, such as, study thesis, are given a 0=No.

QC4. What is the impact of the findings of a study on the overall research being performed?
Since, publication types, such as, book chapter and journal papers, have a greater impact on the overall research being performed, they are considered to be of better quality compared to other types, such as, conference or workshops. Hence, we define the following strategy to score the published articles:

a. Journal publications and book chapters are assigned a score 1.
b. Other officially published articles, such as, conference and workshop are assigned a score 0.5.
c. Remaining unofficially published articles are assigned a score 0.

QC5. Is sufficient information present in the study for clear understanding of readers?
For a clear and adequate explanation about the study, the length of document increases. If the article is short then probably the descriptions provided about the study is brief while detailed explanation requires more number of pages. Hence, for our approach we adopt the following scoring criteria:

a. Number of pages greater than 6, assigned a score 1.
b. Number of pages between 4 and 6, assigned a score 0.5.
c. Number of pages less than 4, assigned a score 0.

QC6. Does the abstract of the paper give a clear idea about the study conducted?
Structured abstracts provide a clear description of the content present in a study. Usually beginning with a brief description about the purpose of the study and background. Followed by an outline of the methodology and summary of results. We assign scores to the studies based on the following:

a. Studies with structured abstracts, given a score 1
b. Studies with clear but unstructured abstracts, given a score 0.5.

QC7. How frequently is the study referred to by other research work?
Studies that provide useful and valid findings are usually cited-by other researchers for their work. Hence, we use the cited-by count of the papers to determine the most reliable ones. We normalize the counts to obtain values between 0 and 1. Assigning the most frequently referred study a score 1 while, papers with least counts have scores closer to 0.

4.2.4 Results

As the result of our approach, we display a list of selected articles with their assigned scores for individual quality criteria as well as their combined final scores. The reviewer can analyze all of the determined scores and according to the importance of each quality criteria, perform the final selection of primary studies.

To simplify the representation of results, we summarize each quality criteria with a specific term and assign the scores for each paper. Table 4.1 represents each quality criteria and their corresponding term:

Quality Criteria:	QC1	QC2	QC3	QC4	QC5	QC6	QC7
Short Terms:	Keyword	Abstract	Author	Pages	Type	Publisher	Cited-by

Table 4.1: Defined quality criteria with corresponding short representations.

4.3 Threat to Validity

For our approach, we realize certain shortcomings that can influence the results provided by the implemented tool. This includes a limitation during the first step, where we define the search terms individually without any consideration of Boolean operators. Since, for every study, the occurrence of each keyword is determined independently, regardless whether they are synonyms or not, this might effect the frequency count analysis to determine the relevancy of papers. For a more accurate content analysis, related terms or synonyms can be grouped together and an occurrence of any one of the terms is treated as a count for the entire group.

Furthermore, to conclude regarding the content relevancy of a paper we only analyze specific parts of the text, that includes, title, abstract and keywords. Although, for an improved content analysis we made an attempt to upload the full text but remained unsuccessful because uploading and parsing of portable document formats (PDFs) is possible but complex task [Laskov and Šrndić, 2011]. Hence, we use only specific parts of studies for the content analysis. Due to this limitation with loading the full text of articles, we also limit quality scoring criteria to the meta data and abstract of articles. For more efficient scoring, quality assessment criteria that analyze complete content of the papers would be beneficial.

Through our approach, we aimed to automate the process of identifying related articles through reference and citation relationships of the selected ones. However, due to certain restrictions imposed by the digital libraries, we could not access the related articles automatically. Hence, for our tool, the user is required to manually identify the referred and citing articles. After identification, the user must upload the meta data and abstracts of these articles to be examined for quality assessment stage.

Even though these limitations exist, our approach proved to be promising since the goal is to support the user in the selection process as much as possible. Some human effort is

still needed to obtain accurate results but the time and effort required for the tasks to be performed entirely manually would surely be reduced using the proposed approach.

4.4 Summary

In this chapter, we described an approach we propose to semi-automate the selection and quality assessment of primary studies for an SLR. We elaborate the fundamental concepts of our approach, including explanation about the input, processing and output of the implemented tool. To initiate our approach, the user uploads search terms and initial set of identified studies. Later on, the identified literature is analyzed based on content relevancy. We suggest to perform backward and forward snowballing to further identify the related articles. Finally, our approach provides the quality scores assigned to each articles based on the set of defined criteria. We concluded the chapter with a discussion regarding some limitations of our approach. To evaluate the validity of results obtained through our approach, we perform a case study that is described in the next chapter.

5. Evaluation

As we described in the previous chapter, our approach must be evaluated to verify its significance for researchers conducting literature reviews. For this, we implement our approach into a web-based application and use a case study for the evaluation. We begin this chapter with a description of the SLR procedure followed by Zahedi et al. [2016], that we select as our case study. Later, we discuss the primary study selection process applied in the case study and repeat the search procedure to obtain results using our approach. To investigate the validity of results obtained through the tool, we perform comparisons with the primary studies selected by Zahedi et al. [2016]. Finally, we discuss some limitations of the implemented web-based application and problems that may arise during the steps.

5.1 Systematic Review Case Study

To explore the possible contribution of our tool for performing more effective systematic reviews, we choose a specific research topic and perform a case study. By evaluating the number of studies correctly selected by our approach we determine its validity. We choose the study performed by Zahedi et al. [2016] for the evaluation. The objective of the chosen SLR was to systematically identify and synthesize knowledge sharing challenges and practices. The reviewers aimed to classify the recurrent challenges and most frequently reported practices in different contextual settings.

Zahedi et al. [2016] performed an SLR for reviewing studies published in the Global Software Development (GSD) literature over a span of years 2000-2014. This study was conducted in September, 2014 following the guidelines published by Kitchenham and Charters [2007] for applying the research method that involves three main phases: defining a review protocol, conducting the review, and reporting the results. The review protocol consisted of these elements: (i) RQs, (ii) search strategy, (iii) inclusion and exclusion criteria, (iv) study selection, (v) study quality assessment, and (vi) data

extraction and synthesis. We use the following information to replicate the study and find the results through our approach:

RQs

The study performed by Zahedi et al. [2016] aimed at gaining an understanding of knowledge sharing challenges, and practices reported by empirical studies in GSD, and identify the contextual settings from which the challenges and practices are found. Thus, the RQs of interest and their respective motivations are briefly described below:

RQ1: What are the knowledge sharing challenges in GSD?
Through the first RQ, reviewers aim to get an overview of the different types of knowledge sharing challenges reported in GSD and identify the challenges that are most frequently reported in the context of GSD.

RQ2: What are the knowledge sharing practices in GSD?
The second RQ focuses on the retrieving appropriate information to gain an understanding of the knowledge sharing practices implemented in GSD.

RQ3: In what contextual settings (i.e.,research methodology and organizational context) are challenges and practices reported?
As a result of the last RQ, the reviewers intend to understand the methodological and organizational contextual settings from which challenges and practices have been reported. Such information can help a reader to better understand the reported research.

Search String

We use the same search strategy as defined by Zahedi et al. [2016] for identifying the relevant literature. The search string is formulated based on the three compartments shown in Table 5.1.

Category	Description
A	"Global Software Development" and all synonyms
B	"Knowledge Management" OR "Knowledge Sharing" OR "Knowledge Transfer" with all synonyms
C	"Challenges" OR "Solutions" with all synonyms

Table 5.1: Structure of search string based on Zahedi et al. [2016].

Based on the keywords, the following search string was constructed by the reviewers:

TITLE-ABS-KEY (("global software development" OR "collaborative software development" OR "global software engineering" OR "distributed software development"

OR “distributed software engineering” OR “offshore software development” OR “offshore software engineering” OR “geographically distributed software development” OR offshor* OR “software outsourcing” OR “software outsource” OR “globally distributed software development” OR “offshore outsourcing” OR “Dispersed teams” OR “distributed teams” OR “virtual teams” OR “globally distributed work” OR “global software teams” OR outsour*)
AND (“knowledge transfer” OR “knowledge shift” OR “knowledge exchange” OR “knowledge distribution” OR “tacit knowledge” OR “explicit knowledge” OR “knowledge transfer process” OR “knowledge flow” OR “organizational knowledge” OR “knowledge acquisition” OR “knowledge management” OR “knowledge creation” OR “knowledge sharing” OR “knowledge retention” OR “knowledge valuation” OR “knowledge use” OR “knowledge application” OR “knowledge discovery” OR “knowledge integration” OR “knowledge theory” OR “knowledge engineering” OR “experience transfer” OR “technology transfer”)
AND (risk* OR challenge* OR tool* OR method* OR problem* OR challeng* OR barrier* OR “best practices” OR model* OR techniq* OR strateg* OR approach* OR process* OR solution* OR obstacle* OR “risk analysis” OR effect* OR “risk factors” OR selection* OR mechanism* OR assessment* OR “evaluation process” OR practice* OR mitigate*))

Data Source

Zahedi et al. [2016] applied the search query on the Scopus indexing system to identify an initial set of potentially relevant studies. The reviewers made this decision based on successful experiences of other researchers with using Scopus [Daneva et al., 2014; Kitchenham et al., 2010]. Burnham [2006] describe some features of Scopus that attract the searchers to use it as a data resource for identifying relevant literature. Scopus facilitates formulating a single and complex search query, as well as track a large number of journals and conferences in CS and SE [Burnham, 2006; Daneva et al., 2014]. Moreover, in their study, Kitchenham et al. [2010] compare results of Scopus with a manual search and observed that Scopus output covered all the relevant papers that used appropriate terminology and published in popular SE literature venues.

Furthermore, Zahedi et al. [2016] selected Scopus as the only means of collecting data because during their pilot search with directly using digital libraries (such as, Springer Link, Wiley Inter Journal Science, IEEEXplore), it was revealed that a number of restrictions are placed by them on large-scale searches on the meta-data of the published papers. Moreover, it was necessary to modify the search string for each single digital library that might result in errors being introduced. Therefore, selecting Scopus helped the reviewers to keep the search string constant while retrieving focused hints that are mostly relevant. Table 5.2 provides the settings that Zahedi et al. [2016] used for applying the search string. In addition to CS, other relevant subject areas, such as, business management, decision science, and economics, are also chosen that may have the potential to publish related studies, especially from the knowledge management perspective.

Digital Library	Scopus
Years	2000-2014
Language	English
Run on	Title/Abstract/Keyword
Subject Areas	"Computer Science", "Business Management and Accounting", "Decision Sciences", "Economics, Econometrics and Finance"
Date of running search string	7th September 2014
Number of Hints	1,320

Table 5.2: Details of the search string applied by Zahedi et al. [2016].

Inclusion and Exclusion Criteria
To select the relevant studies and answer the RQs, the reviewers define certain IC. The inclusion and exclusion criteria were applied to all of the retrieved studies from Scopus database. The criteria defined for including/excluding papers were as follows:

- $C1$ Peer-reviewed papers included only and excluded editorials, position papers, keynotes, reviews, tutorial summaries, panel discussions and short papers (i.e. less than 6 pages).
- $C2$ GSD context—Different collaboration models of GSD such as offshore outsourcing, offshore-insourcing, and near-shore are included.
- $C3$ Included studies reporting empirical research using methods such as case studies, survey, experiments, and ethnographical studies.
- $C4$ The studies based on students' data were included only if they were reporting findings that were expected to contribute to the RQs. If the studies were only from educational perspective (e.g., how to run GSD courses) they were excluded.
- $C5$ Extended papers—If two papers from the same research on the same topic were published in different venues (e.g., a conference and a journal), only the journal paper was included.
- $C6$ All duplicated studies found from different sources were detected and removed.

5.1.1 Case Study: Procedure

The number of studies selected at each stage of the selected case study is shown in Figure 5.1. Based on the applied search settings given in Table 5.2, the search in Scopus returned 1,320 results which the reviewers filtered by reading their title and abstract.

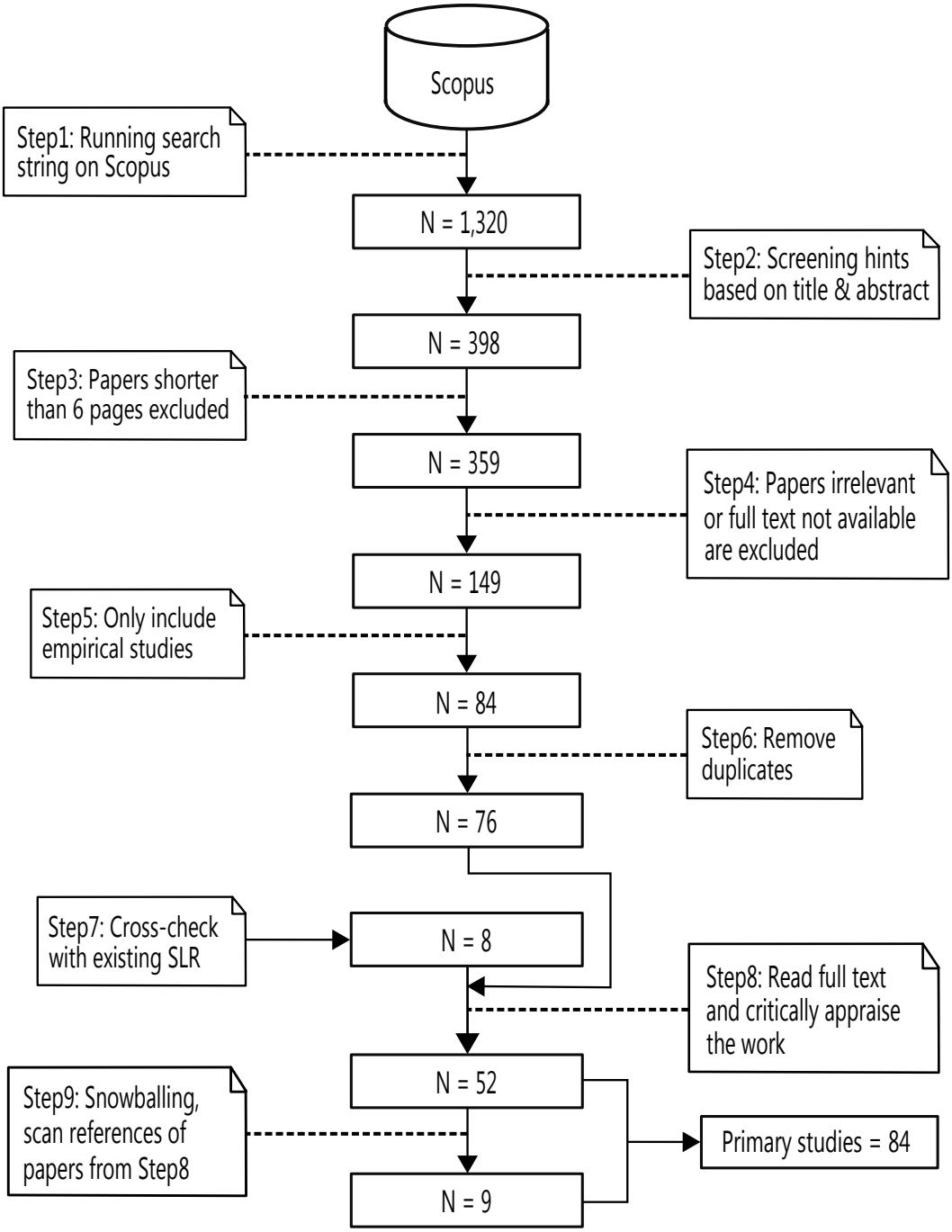

Figure 5.1: Study selection process followed by Zahedi et al. [2016].

The papers that could not be decided upon by reading the titles and abstracts, were retained for the next round of inspection. Articles shorter than 6 pages, duplicates, viewed as irrelevant, or whose full text was not available were excluded. Since, the researchers were interested in empirical studies, papers that were not supported with empirical data were also excluded. Furthermore, those papers that had reported research involving software teams in any type of distributed arrangement were included. At the end of step 6, Zahedi et al. [2016] found 76 papers that met all the IC.

After the list of primary studies was obtained, reviewers performed a crosscheck with the primary studies included in a relevant study conducted previously by Nidhra et al. [2013]. The list of studies found were compared with primary studies included by Nidhra et al. [2013] and 8 studies were added which were either not found by the search or had been excluded based on title/abstract reviews. In the eighth step of selection process, 84 papers (i.e., the total papers selected by Zahedi et al. [2016] and the ones added based on study by Nidhra et al. [2013]) were distributed between two first authors to filter the papers by reading their full text. As a result, 52 studies were selected as primary studies. Then a snowballing technique was followed to scan the references of these 52 selected papers and 181 potentially relevant papers were found. Later, step 2 to step 8 were applied on these papers and finally, 9 of them were included as primary studies. Thus, Zahedi et al. [2016] retrieved 61 primary studies for their study.

5.2 Approach: Collecting Data

For our approach, it was required to provide the initial set of papers obtained from the selected data resource, using the search query defined by Zahedi et al. [2016]. The web-based application expects as input, the initial search results and search terms to process the results. Hence, we obtain the first set of potentially relevant papers from Scopus, using the same search strategy as applied in the case study. Table 5.3 provides the search results we retrieved through our search.

Digital Library	Scopus
Date of running search string	28th September 2017
Number of Hints	1,191

Table 5.3: Details of the search string applied for our approach.

For identifying literature from Scopus, we apply the same refinements as done in the case study. Including, language, year range, subject areas and search categories, as given in Table 5.2. However, number of hints are lesser than that found by Zahedi et al. [2016] for the search. The reason for this could be the difference in execution date of search, there are rapid changes over time in the properties of any digital library due to which there is a possibility to find varying results using the same search query and search refinements [Kitchenham et al., 2015]. Therefore, we continue with this initial set of 1,191 results for further processing and evaluation of our approach.

5.3 Results

As we have discussed previously in Section 4.2, the search terms along with meta data and abstract of identified articles must be provided to the web-based application for processing. Hence, to begin with, we upload the required input, the number of entries inserted in the database were 1,091. We observe that a number of entries could not be stored into the database because of some problems with them. For example, information, such as, author, was missing for some entries or there were duplicate entries. At this point, we decide to overlook such entries as including such results might effect negatively on the quality of our study. Hence, the number of articles to be examined for the next step of the tool were less than the initial number of hints we obtain from Scopus.

Once the meta data and abstract of the articles are loaded, we determine the frequency counts based on titles, abstracts and keywords. We then select the potentially relevant articles by analyzing their titles and keywords frequency counts. Through this first screening, we obtain 455 results from the initial set of 1,091 articles to be processed further. Before proceeding with the next stages of our approach, we remove articles that were shorter than 6 pages, as agreed by Zahedi et al. [2016]. As a result, we obtain 420 results to be examined for the next step. As we mention earlier in Section 4.2.2, the tool expects to receive bibliographic information of references and citations, that must be identified by manually performing forward and backward snowballing for the selected papers. However, performing this step for a large collection of articles demanded a lot of time and due to time constraints in our work we decide to skip this step for future analysis. Thus, to evaluate our approach, we continue with the process to achieve quality scores of papers selected based on content (title, abstract and keyword). We obtain the score for each quality criteria defined previously in Section 4.2.3, the complete set of 420 results we obtain from the tool are given in Section A.2.

5.3.1 Case Study: Results

We retrieve the primary studies selected in the case study from the complete collection of 420 results achieved from the tool, and eventually realize that out of the 52 selected studies, we find only 46 of them. To investigate reasons for these missing results, we further analyze the set of articles in the previous steps and notice that these results are not removed during the screening process instead they were not present in the initial set of papers we load as input. Hence, we overlook these missing entires and evaluate the scores obtained for the identified primary studies selected by Zahedi et al. [2016].

Table 5.4 enlists the papers included in the case study with their corresponding scores assigned through our approach. We sort the results in descending order of the final score for further analysis. The tool provides values for the unique identifiers (P_ID) that represent their entry number in the initial set of papers. To draw valid conclusions we first calculate the average of the final scores of the papers. For the complete collection of 420 results, we determine that:

$$\overline{finalscore} = 3.32$$

Table 5.4: Quality scores of primary studies selected by Zahedi et al. [2016].

P_ID	Title	Quality Criteria Score							
		Keyword	Abstract	Author	Pages	Type	Publisher	Cited-by	Final
152	An empirical study on improving shared understanding of requirements in GSD	0.004	1	0.67	1	1	1	0.01	4.68
44	Knowledge transfer planning and execution in offshore outsourcing: An applied approach	0.028	0.5	1	1	1	1	0.01	4.54
718	Explaining variations in client extra costs between software projects offshored to India	0.025	0.5	0	1	1	1	0.65	4.18
590	Offshore-onsite subgroup dynamics in globally distributed teams	0.023	1	0	1	1	1	0.03	4.05
49	Managing knowledge on communication and information flow in global software projects	0.038	0.5	0.33	1	1	1	0.01	3.88
60	Towards an understanding of enabling process knowing in global software development: A case study	0.045	0.5	0.67	1	1	0.5	0	3.72
989	Embedded knowledge and offshore software development	0.015	0.5	0	1	1	1	0.2	3.72
26	Knowledge sharing for common understanding of technical specifications through artifactual culture	0.031	0.5	0.67	1	1	0.5	0	3.7
955	Knowledge transfer in virtual systems development teams: An exploratory study of four key enablers	0.013	0.5	0	1	1	1	0.18	3.69
160	Spatial knowledge creation and sharing activities in a distributed agile project	0.012	0.5	0.67	1	1	0.5	0	3.68
730	Knowledge transfer and organizational learning in IS offshore sourcing	0.007	0.5	0	1	1	1	0.16	3.67
588	Use of collaborative technologies and knowledge sharing in co-located and distributed teams: Towards the 24-h knowledge factory	0.023	0.5	0	1	1	1	0.11	3.63
23	Knowledge transfer in offshore outsourcing software development projects: An analysis of the challenges and solutions from German clients	0.082	0.5	0	1	1	1	0.04	3.62
725	A US Client's learning from outsourcing IT work offshore	0.023	0.5	0	1	1	1	0.1	3.62
835	Globally distributed component-based software development: An exploratory study of knowledge management and work division	0.018	0.5	0	1	1	1	0.09	3.61
15	Knowledge transfer in IT offshoring relationships: The roles of social capital efficacy and outcome expectations	0.015	0.5	0	1	1	1	0.06	3.58
265	Knowledge sharing practices and the impact of cultural factors: Reflections on two case studies of offshoring in SME	0.022	0.5	0	1	1	1	0.05	3.57
850	Traceability-based knowledge integration in group decision and negotiation activities	0.013	0.5	0	1	1	1	0.06	3.57
132	Knowledge sharing barriers in global teams	0.036	0.5	0	1	1	1	0.02	3.56
34	Integration by communication: Knowledge exchange in global outsourcing of product software development	0.039	0.5	0	1	1	1	0	3.54
621	Seeing remote team members as leaders: A study of US-Scandinavian teams	0.008	0.5	0	1	1	1	0.03	3.54

P_ID	Title	Quality Criteria Score							
		Keyword	Abstract	Author	Pages	Type	Publisher	Cited-by	Final
251	Managing meta-learning in offshore software development environments	0.026	0.5	0	1	1	1	0	3.53
363	"Path to ""Stardom"" in Globally Distributed Hybrid Teams: An Examination of a Knowledge-Centered Perspective using Social Network Analysis"	0.011	0.5	0	1	1	1	0.02	3.53
452	Operational and strategic learning in global software development	0.013	0.5	0	1	1	1	0.02	3.53
498	Exploring the media mix during IT-offshore project	0.033	0.5	0	1	1	1	0	3.53
48	Managing Knowledge in Global Software Development Projects	0.018	0.5	0	1	1	1	0.01	3.53
210	Knowledge management initiatives in offshore software development: Vendors' perspectives	0.018	0.5	0	1	1	1	0	3.52
401	A task-driven approach on agile knowledge transfer	0.019	0.5	0	1	1	1	0	3.52
529	Transactive memory and the transfer of knowledge between onsite and offshore IT outsourcing teams	0.021	0.5	0	1	1	1	0	3.52
52	The knowledge-bridging process in software offshoring from Japan to Vietnam	0.006	0.5	0	1	1	1	0	3.51
711	Mind the gap! Understanding knowledge in global software teams	0.012	0.5	0	1	1	1	0	3.51
762	Collaboration patterns and the impact of distance on awareness in requirements-centred social networks	0.016	0.5	0	1	1	0.5	0.13	3.15
770	Of deadlocks and peopleware - Collaborative work practices in global software development	0.037	0.5	0	1	1	0.5	0.03	3.07
225	Knowledge management in distributed agile software development	0.02	0.5	0	1	1	0.5	0.04	3.06
291	Architectural knowledge management practices in agile global software development	0.026	0.5	0	1	1	0.5	0.02	3.05
316	An understanding of the role of trust in knowledge seeking and acceptance practices in distributed development teams	0.015	0.5	0	1	1	0.5	0.03	3.05
614	Breaching the knowledge transfer blockade in IT offshore outsourcing projects - A case from the financial services industry	0.022	0.5	0	1	1	0.5	0.03	3.05
785	Managing offshore outsourcing of knowledge-intensive projects a people centric approach	0.02	0.5	0	1	1	0.5	0.03	3.05
783	Piloting knowledge transfer in it/is outsourcing relationship-towards sustainable knowledge transfer process learnings from Swiss financial institution	0.028	0.5	0	1	1	0.5	0	3.03
522	Workgroup structures in offshore software development projects: A vendor case study	0.023	0.5	0	1	1	0.5	0	3.02
665	Storytelling - A method to start knowledge transfer in offshore software development teams - research in progress paper	0.022	0.5	0	1	1	0.5	0	3.02
73	The four 'W's of face-to-face - Suggesting an enriched perspective on nearshoring relationship management	0.01	0.5	0	1	1	0.5	0	3.01
324	An experience base with rights management for global software engineering	0.014	0.5	0	1	1	0.5	0	3.01
896	Can distributed software development be agile?	0.038	0.5	0	0	1	1	0.33	2.87
564	Communication, knowledge and co-ordination management in globally distributed software development: Informed by a scientific software engineering case study	0.041	0.5	0	0	1	0.5	0.02	2.06
566	Which groupware tool is the most suitable for this group?	0.014	0.5	0	0	1	0.5	0	2.01

5.3.2 Analysis

In the Venn-diagram shown below, Figure 5.2, we represent the distribution of results provided by the tool. From an entire collection of 420 papers, we observe that 46 primary studies are selected by Zahedi et al. [2016]. Among these selected papers, the tool assigns 15 of them a score below the average final score (false negatives). Furthermore, we also determine that out of the 374 not selected studies, 227 of them have scores higher than the average final score (false positives).

The last 15 papers listed in Table 5.4 have a final score less than the average. We observe that, most of them are conference papers hence, for the case study, type of publication (journal or conference) was not a very significant factor while assessing the studies. Although for our approach, we define low scores for publication types other than journal articles according to QC4, resulting in an overall lower score. We also notice that, few false negatives occur because some papers were included by Zahedi et al. [2016], from a previously conducted study. Therefore, we conclude that the selection process followed by the researchers in the case study is also not perfect, so there can be some deviations in the results.

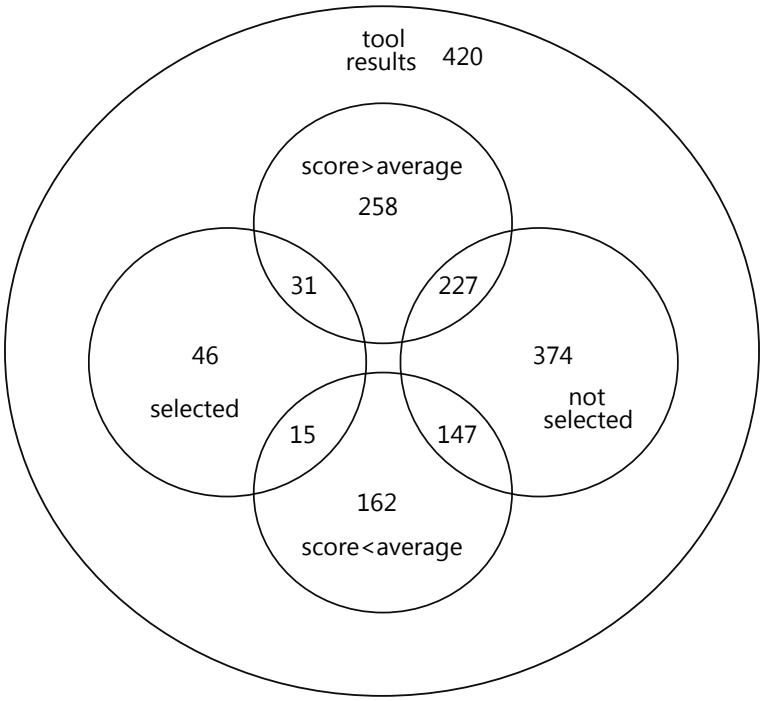

Figure 5.2: Venn-diagram representing results.

Similarly, we also investigate the large number of false positives for our approach. Analyzing the results, we realize that some papers with higher final scores, although have high keyword frequency counts but do not provide direct evidence to answer the defined RQs. As we mentioned earlier that in our approach, we perform the content analysis based on only specific parts of the content and not the entire text hence, reducing the

possibility to retrieve very precise results and increasing the possible occurrence of false positives. On the other hand, we also notice that some papers fulfill the content criteria but do not completely satisfy the criteria defined by Zahedi et al. [2016] for inclusion. Further examining the results, we recognize that a full text analysis is specifically significant to determine regarding the IC C3 and C4, listed in the first section of this chapter.

Moreover, we also realize that the implemented tool is not perfect in every aspect at the moment, so there are possible variations in the results. We determine quality scores of the papers, based on a limited set of assessment criteria analyzing only the meta data and abstracts. However, as quality assessment criteria are individually defined by researchers conducting an SLR, we must consider the fact that reviewers might assess the quality of papers in a manner different from our approach. For instance, in the case study we select for evaluating our approach, the researchers do not specify the quality assessment criteria applied to analyze the papers. Hence, we are unable to define the significance of individual quality scoring criteria considered in our approach. For example, we obtain some results with low keyword counts satisfy the quality scoring criteria (QC2 - QC7) and obtain an overall higher score. Therefore, we consider these limitations the reason for deviation in results obtained from the tool.

5.3.3 Discussion of Results

To determine the performance of our approach, we use the *confusion matrix* as shown in Table 5.5. We summarize the results in this form to measure the accuracy, recall and precision of our approach. These are the most commonly used performance measures to verify the efficiency of a document retrieval approach. We compute each of them for our approach, further in this section.

True Positives (TP) = 31	False Positives (FP) = 227
False Negatives (FN) = 15	True Negatives = 147

Table 5.5: Confusion matrix.

Accuracy
To determine the closeness of results obtained from the tool with those of the case study, we measure the accuracy using Equation 5.1. We achieve 42.3% accuracy for our approach, that means based on the entire data collection of 420 papers, 178 of them have been correctly categorized.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.1}$$

$$accuracy_{\%} = 42.3\%$$

Recall
To measure the number of positives retrieved using the approach, we calculate recall using Equation 5.2. Recall is also referred to as the true positive rate, for determining the correct papers retrieved successfully. For our approach, the recall is calculated to be 67.4%.

$$recall = \frac{TP}{TP + FN} \tag{5.2}$$

$$recall_{\%} = 67.4\%$$

Precision
To conclude regarding the number of correct results from the entire collection, we calculate the precision using Equation 5.3. Due to a large number of false positives, the precision of our approach is found to be 12.0%.

$$precision = \frac{TP}{TP + FP} \tag{5.3}$$

$$precision_{\%} = 12.0\%$$

Through our experimental evaluation, we observe that although the manual effort for performing selection and quality assessment of articles is reduced, the precision of our approach is 12.0% with an accuracy and recall of, 42.3% and 67.4% respectively. The high number of false positives obtained through the tool are mainly due to the fact that, our approach performs analysis using only meta data and abstract. It is most likely that, if the full-text of papers is analyzed, the results obtained will be more precise. However, the overall calculated accuracy and recall indicate that our approach improves the article selection step, reducing the required amount of time and effort.

5.4 Threat to Validity

In this section, we discuss some limitations of the study performed. Firstly, the case study we select for evaluating our approach is conducted by researchers working in the specific field. While, we replicate the study with limited expertise and previous knowledge about the topic. Hence, there are possibilities of errors while screening of results based on titles and keyword counts. Furthermore, we evaluate the approach using only a single data resource, but analyzing a larger dataset from different resources would be beneficial.

The deviations of results, found from our approach, with the case study is also because we analyze only the meta data and abstract of papers. An analysis of complete content would most likely result in higher precision and efficiency. Further, we also skipped the

step of manually identifying the related articles and uploading them. As a result, a reduced number of primary studies are retrieved using the tool (i.e., 46 out of 52).

Additionally, there are some limitations to the quality scoring procedure we follow. As we mention earlier, the lack of a standard procedure for quality assessment of papers urge the researchers to define their own set of criteria according to the topic under observation. For the selected case study, the reviewers did not specify the criteria followed to assess the quality of results. Hence, the set of quality criteria defined for our approach may differ from the preferences of reviewers while conducting the study.

5.5 Summary

In this chapter, we performed the evaluation of our approach using the study conducted by Zahedi et al. [2016]. We began with defining the objectives and search strategy followed in the case study. Later, we explained the procedure followed to replicate the study and the results obtained from the tool. We analyzed the results retrieved from our approach with the primary studies of the case study to draw meaningful conclusions. We concluded the chapter with a discussion about some limitations of our study.

6. Conclusion

SLRs are one of the most important techniques to synthesize relevant evidence related to a particular research area or topic [Kitchenham and Charters, 2007]. During the SLR process, selection of primary studies is the most crucial stage, which if performed manually, can be very time consuming. Hence, SE researchers are interested in applying an approach to semi-automate this step [Hassler et al., 2016]. In this work, we begin with analyzing the existing approaches proposed by researchers of the CS domain to aid reviewers during the selection of primary studies. Based on the findings, we conclude that there is some research conducted in this regard but quality scoring of articles is still usually performed manually. Therefore, we propose an approach to efficiently retrieve most promising primary studies to answer the defined RQs.

Through our approach, we aim to reduce the effort and time required to identify the most relevant results for an SLR. The method we proposed is based on text mining techniques, specifically, information extraction, term recognition and categorization. In our approach, the selection is performed based on the content analysis. Presently, we consider inly the title, abstract and keywords of the papers to determine their relevancy. Further, to identify related articles our approach uses the citation links of papers, including both references and cited-by papers. The last stage of our approach, involved quality scoring of the papers based on certain criteria. Once the final scores are assigned, the list of articles with their corresponding scores is displayed to the user.

To evaluate the proposed approach, we selected a case study performed by Zahedi et al. [2016]. We execute the initial stages of the SLR process similar to the procedure followed in the study. Later on, instead of performing the selection process manually we use a web-based application implemented based on our approach. After obtaining the results, we determined the efficiency of our approach using the performance measures. The accuracy of our approach is calculated to be 42.3% with a recall of 67.4%. Based on these results, we think that our approach proves to be suitable for supporting the selection of primary studies for an SLR. However, to improve the validity of results,

we recommend some future improvements in the next chapter. We concluded our work with an explanation of the limitations of our approach and suggest procedures that could be followed to overcome them.

7. Future Work

During our work, we identify several aspects of our approach that require further improvements in future. The most significant requirement for improving the precision of results is to consider the full-text of papers for content analysis and quality scoring. We believe that, by incorporating this in our approach, the selection and quality assessment step during the SLR process could be performed more efficiently and also improve the reliability of obtained results.

Currently, our approach treats the defined search terms individually without consideration of the Boolean operators. To initiate the process, the user provides each term separately regardless of being related or synonyms. For a more effective procedure, the related terms must be grouped to determine the frequency counts. As a result, if any of the synonyms are found in the content, frequency count for the entire group will be increased.

Furthermore, we suggest to incorporate a method to semi-automate the step of identifying the papers related to the selected ones (i.e., citations and references). At present, our approach requires to perform this step manually which is quite time consuming. Apart from that, for evaluation of our approach the selected case study considers Scopus as the only data source. However, for further investigation more data resources can be used to examine a larger set of initial results.

We also suggest to improve the quality scoring procedure proposed in our approach. For the present work, we define some specific quality criteria based on meta data and abstract of papers, although researchers usually assess articles based on criteria appropriate to the study being conducted. Hence, further research must be performed to determine the factors considered most important by researchers for assessing studies, based on which a significance factor can be assigned to each of the criteria defined in our approach. In this way, it is possible to obtain more accurate results with reduced amount of manual effort.

A. Appendix

A.1 Search results of the SLR.

Title	Year
SESRA: A Web-based Automated Tool to Support the Systematic Literature Review Process	2015
A Visual Analysis Approach to Update Systematic Reviews	2014
Characterizing Software Requirements Elicitation Processes: A Systematic Literature Review	2016
A Survey on Aims and Environments of Diversification and Obfuscation in Software Security	2016
A Survey of Formal Methods in Self-adaptive Systems	2012
Usability of Requirements Techniques: A Systematic Literature Review	2016
A Systematically Conducted Literature Review: Quality Attribute Variability in Software Product Lines	2012
A Systematic Literature Review of Traceability Approaches Between Software Architecture and Source Code	2014
Energy-Efficient Networking Solutions in Cloud-Based Environments: A Systematic Literature Review	2015
Environment Modeling in Model-based Testing: Concepts, Prospects and Research Challenges: A Systematic Literature Review	2015
Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering	2014
A Systematic Literature Review on the Description of Software Architectures for Systems of Systems	2015
The Adoption of Capture-recapture in Software Engineering: A Systematic Literature Review	2015
Evidence in Software Architecture, a Systematic Literature Review	2013
The Systematic Review of Literature in LIS: An Approach	2016
A Systematic Review of Web Resource Estimation	2012
Systematic Literature Review: Ingenious Software Project Management While Narrowing the Impact Aspect	2016
A Method to Support Search String Building in Systematic Literature Reviews Through Visual Text Mining	2015
Fusion of Face and Gait for Biometric Recognition: Systematic Literature Review	2016
A Reviewbased Comparative Study of Bad Smell Detection Tools	2016
Systematic Software Process Development: Where Do We Stand Today?	2013
Protocol for a Systematic Literature Review of Research on the Wikipedia	2009
Collaborative Exchange of Systematic Literature Review Results: The Case of Empirical Software Engineering	2015
Systematic Literature Studies: Database Searches vs. Backward Snowballing	2012
A Conceptual Model to Address Threats to Validity in Controlled Experiments	2013
Is There a Place for Qualitative Studies when Identifying Effort Predictors?: A Case in Web Effort Estimation	2014
Applicability of the Semiotic Inspection Method: A Systematic Literature Review	2011
A Tertiary Study: Experiences of Conducting Systematic Literature Reviews in Software Engineering	2013
Does Pareto's Law Apply to Evidence Distribution in Software Engineering? An Initial Report	2014
Effort Estimation in Agile Software Development: A Systematic Literature Review	2014
Using Qualitative Metasummary to Synthesize Empirical Findings in Literature Reviews	2014
Accessibility in Rich Internet Applications: People and Research	2012
Software Paradigms, Assessment Types and Non-functional Requirements in Model-based Integration Testing: A Systematic Literature Review	2014
Is There a Future for Empirical Software Engineering?	2016
Research Challenges in Business Process Adaptability	2014
Tools to Support Systematic Reviews in Software Engineering: A Cross-domain Survey Using Semi-structured Interviews	2015
Self-adaptation for Cyber-physical Systems: A Systematic Literature Review	2016
A Systematic Review of Design Diversity-based Solutions for Fault-tolerant SOAs	2013
Understanding Learning at a Glance: An Overview of Learning Dashboard Studies	2016
Factors Influencing Hybrid Self-Regulated and Collaborative Learning for End-User Training: A Systematic Literature Review	2013
Feature Weighting Techniques for CBR in Software Effort Estimation Studies: A Review and Empirical Evaluation	2014
A Systematic Review on Mining Techniques for Crosscutting Concerns	2013
Challenges and Recommendations for the Design and Conduct of Global Software Engineering Courses: A Systematic Review	2015

Title	Year
A Systematic Review of System-of-systems Architecture Research	2013
A Systematic Review on Metamodels to Support Product Line Architecture Design	2016
Cloud-based Bibliometric Analysis Service for Systematic Mapping Studies	2015
Capturing Cost Avoidance Through Reuse: Systematic Literature Review and Industrial Evaluation	2016
A Systematic Review of Business and Information Technology Alignment	2013
Cross- vs. Within-company Cost Estimation Studies Revisited: An Extended Systematic Review	2014
'Follow the Moon' Development: Writing a Systematic Literature Review on Global Software Engineering Education	2015
Using Rule-based Classifiers in Systematic Reviews: A Semantic Class Association Rules Approach	2015
User Involvement in Software Development and System Success: A Systematic Literature Review	2013
A Literature Review of Agile Practices and Their Effects in Scientific Software Development	2011
A Critical Analysis of Studies That Address the Use of Text Mining for Citation Screening in Systematic Reviews	2016
Sustained Agile Usage: A Systematic Literature Review	2013
Extended Boolean Retrieval for Systematic Biomedical Reviews	2010
Agile Practices in Practice: A Mapping Study	2014
Approximation of Statistical Implicative Analysis to Learning Analytics: A Systematic Review	2016
Using Grammars for Pattern Recognition in Images: A Systematic Review	2013
Engineering Quality and Reliability in Technology-Assisted Review	2016
Experiences from Using Snowballing and Database Searches in Systematic Literature Studies	2015
Representation of Gestures in Syntactic Analysis: A Systematic Literature Review	2016
Changing Approaches to Research Synthesis Affect Social and Intellectual Structures of Science	2016
On the Use of Security Principles and Practices for Architecting Cyber-physical Systems	2016
PaperQuest: A Visualization Tool to Support Literature Review	2016
Model-based Testing in Industry: A Case Study with Two MBT Tools	2010
Evidence-based Medicine, the Essential Role of Systematic Reviews, and the Need for Automated Text Mining Tools	2010
Reviewing the Flipped Classroom Research: Reflections for Computer Science Education	2014
Ten Years of Rich Internet Applications: A Systematic Mapping Study, and Beyond	2014
A Life-cycle Perspective on Online Community Success	2009
Application of IT in Healthcare: A Systematic Review	2016
Data Quality in Empirical Software Engineering: A Targeted Review	2013
The Development of Social Network Analysis Research in Mainland China: A Literature Review Perspective	2011
Characterisation of Challenges for Testing of Adaptive Systems	2016
Sustainability Evaluation of Software Architectures: A Systematic Review	2011
A Systematic Mapping Study on Serious Game Quality	2014
Big Data - Characterizing an Emerging Research Field Using Topic Models	2014
Empirical Studies of Embedded Software Development Using Agile Methods: A Systematic Review	2012
Knowledge Extraction for Literature Review	2016
Requirements Elicitation from Business Process Model in BPMN: A Systematic Review	2016
Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies	2015
Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience	2011
Characterizing DevOps by Hearing Multiple Voices	2016
The Impacts of Software Process Improvement on Developers: A Systematic Review	2012
A Review of Continuous Authentication Using Behavioral Biometrics	2016
The Role of ICT in Addressing the Challenges of Age-related Falls: A Research Agenda Based on a Systematic Mapping of the Literature	2015
Concurrent Software Testing in Practice: A Catalog of Tools	2015
TrueView: Harnessing the Power of Multiple Review Sites	2015
Software Fault Prediction for Object Oriented Systems: A Literature Review	2011
Claims and Supporting Evidence for Self-adaptive Systems: A Literature Study	2012
A Systematic Survey of Self-Protecting Software Systems	2014
UML Consistency Rules: A Systematic Mapping Study	2014
Concerns in Software Development: A Systematic Mapping Study	2014
Software Requirements Patterns: A State of the Art and the Practice	2015
Systematic Mapping Study on High-level Language Virtual Machines	2010
Assessing Review Recommendation Techniques Under a Ranking Perspective	2016
Fault-localization Techniques for Software Systems: A Literature Review	2014
A Methodology for Evaluation of Web-based Scholarship	2008
Effective Assignment and Assistance to Software Developers and Reviewers	2016
Approaches for Prioritizing Feature Improvements Extracted from App Reviews	2016
Support Mechanisms to Conduct Empirical Studies in Software Engineering: A Systematic Mapping Study	2015
A Review of Research on Risk Analysis Methods for IT Systems	2013
Two Peers Are Better Than One: Aggregating Peer Reviews for Computing Assignments is Surprisingly Accurate	2009
Analysis and Testing of Matlab Simulink Models: A Systematic Mapping Study	2013
A Review of Literature in Modeling Approaches for Sustainable Development	2013
Identify Online Store Review Spammers via Social Review Graph	2012
Annotated Program Examples As First Class Objects in an Educational Digital Library	2008
Analyzing the Quality of Information Solicited from Targeted Strangers on Social Media	2013
Reviewing Technical Approaches for Sharing and Preservation of Experimental Data	2014
The Demonstration of the Reviewer's Assistant	2012
The Impact of Cross-platform Development Approaches for Mobile Applications from the User's Perspective	2016
A Mapping Study on Bayesian Networks for Software Quality Prediction	2014
Review Spam Detection via Temporal Pattern Discovery	2012
Retrieval of Relevant Opinion Sentences for New Products	2015
Review Spam Detection via Time Series Pattern Discovery	2012
A Systematic Approach Towards User-centric Privacy and Security for Smart Camera Networks	2010
Shepherding the Crowd Yields Better Work	2012
Comparing Pre Commit Reviews and Post Commit Reviews Using Process Simulation	2016
A Systematic Approach for Performance Evaluation Using Process Mining: The POSIDONIA Operations Case Study	2016

Title	Year
Information-seeking Behaviors of Computer Scientists: Challenges for Electronic Literature Search Tools	2013
Trust in e-Participation: Literature Review and Emerging Research Needs	2014
TheAdvisor: A Webservice for Academic Recommendation	2013
Process Simulation for Software Engineering Education	2015
Towards Methods for the Collective Gathering and Quality Control of Relevance Assessments	2009
'Helpfulness' in Online Communities: A Measure of Message Quality	2009
Improving Code Review Effectiveness Through Reviewer Recommendations	2014
Risk Areas in Embedded Software Industry Projects	2010
Show Me Your Eyes! The Combined Use of Eye Tracking and Virtual Reality Applications for Cognitive Assessment	2015
Building Reputation and Trust Using Federated Search and Opinion Mining	2012
Mining Complaints to Improve a Product: A Study About Problem Phrase Extraction from User Reviews	2016
Data Sets and Data Quality in Software Engineering: Eight Years On	2016
Can You Certify Your Software to MC/DC?: A Static Analysis Approach to Account for the Number Test Cases	2016
Needs-based Analysis of Online Customer Reviews	2007
A Systematic Mapping Study on Mining Software Repositories	2016
A Systematic Survey on the Design of Self-adaptive Software Systems Using Control Engineering Approaches	2012
Measuring and Fingerprinting Click-spam in Ad Networks	2012
A Literature Review About Technology Startups Ecosystems	2016
What Did Authors Value in the CHI'16 Reviews They Received?	2016
Research Paper Recommender System Evaluation: A Quantitative Literature Survey	2013
Social Network Analysis in Open Source Software Peer Review	2014
CMAP: Effective Fusion of Quality and Relevance for Multi-criteria Recommendation	2011
Supporting Customers' Decision Making with Rated Tags	2014
Accept, Decline, Postpone: How Newcomer Productivity is Reduced in English Wikipedia by Pre-publication Review	2014
A Literature Review on Enterprise Social Media Collaboration in Virtual Teams: Challenges, Determinants, Implications and Impacts	2016
On the Alignment of Source Code Quality Perspectives Through Experimentation: An Industrial Case	2015
Testing Within Constraints: A Practical Approach	2016
A Robust Model for Paper Reviewer Assignment	2014
Beyond Keyword Search: Discovering Relevant Scientific Literature	2011
Hunting for Truly Relevant Articles in Bioinformatics Literature: A Preliminary Study	2010
User Experience Dimensions: A Systematic Approach to Experiential Qualities for Evaluating Information Interaction in Museums	2016
Sensor Network Data Fault Types	2009
Deepening the Role of the User: Neuro-Physiological Evidence As a Basis for Studying and Improving Search	2016
The Promises and Perils of Mining GitHub	2014
Twitter Zombie: Architecture for Capturing, Socially Transforming and Analyzing the Twittersphere	2012
Search Engine Overlaps: Do They Agree or Disagree?	2007
Relevance Criteria for e-Commerce: A Crowdsourcing-based Experimental Analysis	2009
Reviewing Children's Collaboration Practices in Storytelling Environments	2011
A Preliminary Mapping Study of Approaches Bridging Software Product Lines and Service-oriented Architectures	2011
Evaluation of Machine-learning Protocols for Technology-assisted Review in Electronic Discovery	2014
Automating Risk of Bias Assessment for Clinical Trials	2014
A Boosted SVM Based Ensemble Classifier for Sentiment Analysis of Online Reviews	2013
Inferring Undiscovered Public Knowledge by Using Text Mining-driven Graph Model	2014
Characterizing Web-Based Tutorials: Exploring Quality, Community, and Showcasing Strategies	2014
A Generic Approach for Systematic Analysis of Sports Videos	2012
Cloud Computing Adoption Assessment Model (CAAM)	2011
Extracting Customer Knowledge from Online Consumer Reviews: A Collaborative-filtering-based Opinion Sentence Identification Approach	2009
Persuasive System Design: State of the Art and Future Directions	2009
Discovery Tools for Open Access Repositories: A Literature Mapping	2016
Social Haystack: Dynamic Quality Assessment of Citizen-Generated Content During Emergencies	2015
Discriminating Graph Pattern Mining from Gene Expression Data	2016
Gradual Transition Detection in Historic Film Material & Mdash: A Systematic Study	2011
A Methodology to Evaluate Important Dimensions of Information Quality in Systems	2015
Evaluating Link-based Recommendations for Wikipedia	2016
Modeling Adaptation with Klaim	2012
Experiences with Using the Systematic Method for Architecture Recovery (SyMAR)	2013
Performance Analysis of SVM Ensemble Methods for Air Pollution Data	2016
A Preliminary Structure of Software Outsourcing Vendors' Readiness Model	2010
A Qualitative Review on 3D Coarse Registration Methods	2015
Process Mining for Clinical Processes: A Comparative Analysis of Four Australian Hospitals	2015
Process Performance Management: Illuminating Design Issues Through a Systematic Problem Analysis	2011
Effort Estimation in Agile Software Development: A Survey on the State of the Practice	2015
Does Lean Imply Green?: A Study of the Power Performance Implications of Java Runtime Bloat	2012
Does Lean Imply Green?: A Study of the Power Performance Implications of Java Runtime Bloat	2012
Making Usability Recommendations Useful and Usable	2007
Architectural Flexibility in a Software-system's Life-cycle: Systematic Construction and Exploitation of Flexibility	2012
A Systematic Assessment of Smartphone Usage Gaps	2016
Data Sets and Data Quality in Software Engineering	2008
Case Study Research to Reflect Societal and Ethical Issues: Introduction of a Research Implementation Plan for ICTs	2016
Semantics-Based Cross-domain Collaboration Recommendation in the Life Sciences: Preliminary Results	2015
Understanding Online Reviews: Funny, Cool or Useful?	2015
Systematic Intermediate Sequence Removal for Reduced Memory Accesses	2007
Organize, Socialize, Benefit: How Social Media Applications Impact Enterprise Success and Performance	2013
A Review of Difference Score Research in the is Discipline with an Application to Understanding the Expectations and Job Experiences of It Professionals	2011

Title	Year
Where to Place Your Next Restaurant?: Optimal Restaurant Placement via Leveraging User-Generated Reviews	2016
Web Page Classification: Features and Algorithms	2009
Peer Grading in a Course on Algorithms and Data Structures: Machine Learning Algorithms Do Not Improve over Simple Baselines	2016
AxGames: Towards Crowdsourcing Quality Target Determination in Approximate Computing	2016
Usability Recommendations in the Design of Mixed Interactive Systems	2009
Can People Collaborate to Improve the Relevance of Search Results?	2008
Locality Analysis: A Nonillion Time Window Problem	2014
The Perception of Egocentric Distances in Virtual Environments - A Review	2013
Assessments in Global Software Development: A Tailorable Framework for Industrial Projects	2010
A Real-Time Eye Tracking Based Query Expansion Approach via Latent Topic Modeling	2015
Decision-analytic Models for Breast Cancer: Do Currently Published Models Meet the Requirements of Personalized Medicine?	2008
Digital Curation and Costs: Approaches and Perceptions	2016
Inner Source Definition, Benefits, and Challenges	2016
Discrete-event Simulation Optimization: A Review of Past Approaches and Propositions for Future Direction	2013
A Break in the Clouds: Towards a Cloud Definition	2008
How Assessing Plasticity Design Choices Can Improve UI Quality: A Case Study	2010
A Usability Scale for Handheld Augmented Reality	2014
Which Vertical Search Engines Are Relevant?	2013
Complex Objects Ranking: A Relational Data Mining Approach	2010
When to Stop Reviewing Documents in eDiscovery Cases: The Lit I View Quality Monitor and Endpoint Detector	2013
A Catalog of Architectural Tactics for Cyber-Foraging	2015
A Survey on Cloud Interoperability: Taxonomies, Standards, and Practice	2013
Objective Safety Compliance Checks for Source Code	2014
Quality, Trust, and Utility of Scientific Data on the Web: Towards a Joint Model	2011
Trust in Health Websites: A Review of an Emerging Field	2010
Exploiting User Feedback to Learn to Rank Answers in Q&A Forums: A Case Study with Stack Overflow	2013
Reducing Human Effort and Improving Quality in Peer Code Reviews Using Automatic Static Analysis and Reviewer Recommendation	2013
What is Translational Literature and How to Classify It?: Crowd-sourcing As a Starting Point for Corpus Building and Type Distinction in Comparative Literature	2016
An Enhanced Quality of Context Evaluating Approach in the e-Health Sensor Platform	2015
An Analysis Framework for Ontology Querying Tools	2013
Issues on Quality Assessment of SNOMED CT Subsets: Term Validation and Term Extraction	2009
Citation Recommendation Without Author Supervision	2011
A First Measurement Look at the Deployment and Evolution of Thelocator/Id Separation Protocol	2013
Randomized Load Balancing with General Service Time Distributions	2010
Cognitive Radio Networks: Realistic or Not?	2013
Optimising Multiple Quality Criteria of Service-oriented Software Architectures	2009
Towards a User Experience in Immersive Virtual Environment Model: A Review	2015
Mitigating Primary User Emulation Attacks in Dynamic Spectrum Access Networks Using Hypothesis Testing	2009
jMarkov Package: A Stochastic Modeling Tool	2012
Supervised, Unsupervised, and Semi-Supervised Feature Selection: A Review on Gene Selection	2016
Developing Quality Scores for Electronic Health Records for Clinical Research: A Study Using the General Practice Research Database	2011
Industrial Evaluation of the Impact of Quality-driven Release Planning	2014
Impact of Review-Set Selection on Human Assessment for Text Classification	2016
Incorporating Occupancy into Frequent Pattern Mining for High Quality Pattern Recommendation	2012
A Neural Network Based Model for VoIP Speech Quality Prediction	2009
MinePhos: A Literature Mining System for Protein Phoshphorylation Information Extraction	2012
Feature Space Optimization for Content-based Image Retrieval	2012
On the Prioritization of Data Quality Challenges in e-Health Systems in South Africa	2015
A Review of Accessibility Requirements in Elderly Users' Interactions with Web Applications	2012
Performance Evaluation of Relevance Feedback Methods	2008
Towards Automatic Thematic Sheets Based on Discursive Categories in Biomedical Literature	2011
Mining Product Intention Rules from Transaction Logs of an Ecommerce Portal	2009
SAMBA: A Semi-automatic Method for Measuring Barriers of Accessibility	2007
OncoViz: A User-centric Mining and Visualization Tool for Cancer-related Literature	2010
Learning Personal + Social Latent Factor Model for Social Recommendation	2012
Why is Dynamic Analysis Not Used As Extensively As Static Analysis: An Industrial Study	2014
Social Interaction and Reflection for Behaviour Change	2014
A Comparative Approach to Enhance Information Interaction Design of Visual Analytics Systems	2014
Merging Multiple Criteria to Identify Suspicious Reviews	2010
Beyond Lists: Studying the Effect of Different Recommendation Visualizations	2012
ReQ-ReC: High Recall Retrieval with Query Pooling and Interactive Classification	2014
Integration of Psycho-social Models and Methods in NATO's Approach to Operations: a Review of NATO Research and Technology Organization (RTO) Systems Analysis Studies (SAS - 074)	2009
Quality Assessment of Multiobjective Optimisation Algorithms in Component Deployment	2009
Evaluating New Search Engine Configurations with Pre-existing Judgments and Clicks	2011
Quality Factors for Dynamic Evolution in Composition-based Distributed Applications	2011
Competing or Aiming to Be Average?: Normification As a Means of Engaging Digital Volunteers	2014
Trustworthiness Criteria for Supporting Users to Assess the Credibility of Web Information	2013
A Taxonomy of Functional Units for Information Use of Scholarly Journal Articles	2010
Development and Usage of Chinese Medicine Supporting System Based on Post-Controlled Machinery	2009
Fast Top-k Retrieval for Model Based Recommendation	2012
Viewing by Interactions: Media-oriented Operators for Reviewing Recorded Sessions on Tv	2011

Title	Year
Application Case Study of Machine Learning Techniques Towards a Fault Diagnosis System for a Manufacturing Plant Environment	2012
Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects	2015
A Review of Wireless-Sensor-Network-Enabled Building Energy Management Systems	2014
Towards the Assessment of Semantic Similarity Analysis of Protein Data: Main Approaches and Issues	2012
Continual Monitoring of Code Quality	2011
The Flipped Classroom Method: Lessons Learned from Flipping Two Programming Courses	2016
Guaranteeing End-to-end Quality-of-service with a Generic Routing Approach	2014
Speeding Up Glauber Dynamics for Random Generation of Independent Sets	2015
Speeding Up Glauber Dynamics for Random Generation of Independent Sets	2015
The Restoration of Low-quality Audio Recordings Based on Non-negative Matrix Factorization and Perceptual Assessment by Means of the Ebu Mushra Test Method	2010
Recommending Judgment Targets for Rating Provision	2014
Software Systems Through Complex Networks Science: Review, Analysis and Applications	2012
Computational Approach to Biological Validation of Protein-protein Interactions Discovered Using Literature Mining	2008
Human Behaviour Consistent Relevance Feedback Model for Image Retrieval	2007
A Literature-based Method to Automatically Detect Learning Styles in Learning Management Systems	2012
Jointly Modeling Aspects, Ratings and Sentiments for Movie Recommendation (JMARS)	2014
Unified Relevance Models for Rating Prediction in Collaborative Filtering	2008
Mining Knowledge from Databases: An Information Network Analysis Approach	2010
Creating Effective Decision Aids for Complex Tasks	2008
Offline Evaluation and Optimization for Interactive Systems	2015
DiMaC: A Disguised Missing Data Cleaning Tool	2008
Towards a Model to Support in Silico Studies of Software Evolution	2012
Predicting Link Quality Using Supervised Learning in Wireless Sensor Networks	2007
Failure Rates in Introductory Programming Revisited	2014
Understanding and Characterizing Insights: How Do People Gain Insights Using Information Visualization?	2008
Markov Chains with Block-structured State-dependent Events: Formulation and Correlation Analysis	2010
Predictive Rule Discovery from Electronic Health Records	2010
Teaching Code Review Management Using Branch Based Workflows	2016
Towards Distributed Storage Resource Management Using Flow Control	2008
Re-ranking Bibliographic Records for Personalized Library Search	2012
On the Quality of Semantic Interest Profiles for Onine Social Network Consumers	2016
Quantifying QoS Requirements of Network Services: A Cheat-proof Framework	2011
Combining User Reputation and Provenance Analysis for Trust Assessment	2016
Distinguishing Citation Quality for Journal Impact Assessment	2009
Ranking with Tagging As Quality Indicators	2008
Modelling Software Engineering Research with RSML	2014
Analyzing Confidentiality and Privacy Concerns: Insights from Android Issue Logs	2015
The Application of Recommender Systems in a Multi Site, Multi Domain Environment	2015
Research Topic Recommendation Service: An Approach Based on Semantic Relation in Context	2008
Aspect-driven, Data-reflective and Context-aware User Interfaces Design	2013
DiMaC: A System for Cleaning Disguised Missing Data	2008
Multi-angle View on Preference Elicitation for Negotiation Support Systems	2009
Emotional Fingerprint from Authors in Classical Literature	2016
Boosting the Ranking Function Learning Process Using Clustering	2008
Multi-relational Matrix Factorization Using Bayesian Personalized Ranking for Social Network Data	2012
Deploying an Interactive Machine Learning System in an Evidence-based Practice Center: Abstrackr	2012
Enterprise Architecture Benefit Realization: Review of the Models and a Case Study of a Public Organization	2016
The State of the Art in Content-based Image Retrieval in P2P Networks	2010
Seed-detective: A Novel Clustering Technique Using High Quality Seed for K-means on Categorical and Numerical Attributes	2011
Learning Global Term Weights for Content-based Recommender Systems	2016
The Tortoise and the (Soft)Ware: Moore's Law, Amdahl's Law, and Performance Trends for Human-machine Systems	2014
A Comparative Study of Mixed Mode Simulation Methods: Approaches in Research and Practice	2015
A Survey on Assessment and Ranking Methodologies for User-Generated Content on the Web	2015
Using Online Controlled Experiments to Examine Authority Effects on User Behavior in Email Campaigns	2016
Categorical Proportional Difference: A Feature Selection Method for Text Categorization	2008
Discovering Air Quality Patterns in Urban Environments	2016
Content-based Trust and Bias Classification via Biclustering	2012
A Practical Framework for Ethics: The PD-net Approach to Supporting Ethics Compliance in Public Display Studies	2013
Coverage is Not Strongly Correlated with Test Suite Effectiveness	2014
DSMLs for Enterprise Architecture Management: Review of Selected Approaches	2012
Challenges for Usability Testing in Ubiquitous Systems	2014
Predicting Handoffs in 3G Networks	2012
Detecting Reviewer Bias Through Web-based Association Mining	2008
Creativity Factor Evaluation: Towards a Standardized Survey Metric for Creativity Support	2009
Traffic-aware Channel Assignment in Wireless LANs	2007
Utility Maximization in Peer-to-peer Systems	2008
Utility Maximization in Peer-to-peer Systems	2008
Exploiting Query Reformulations for Web Search Result Diversification	2010
Comparison of Decision-Making Strategies for Self-Optimization in Autonomic Computing Systems	2012
Insider Threat Assessment: A Model-Based Methodology	2014
Structured Correspondence Topic Models for Mining Captioned Figures in Biological Literature	2009
Integration of Literature with Heterogeneous Information for Genes Correlation Scoring	2013
Utilizing Review Summarization in a Spoken Recommendation System	2010
Review Sentiment Scoring via a Parse-and-paraphrase Paradigm	2009
Entity Linking for Biomedical Literature	2014

Title	Year
Cloud-enabled Privacy-preserving Collaborative Learning for Mobile Sensing	2012
Automated Hypothesis Generation Based on Mining Scientific Literature	2014
Threats on Building Models from CVS and Bugzilla Repositories: The Mozilla Case Study	2007
Analysis of Hot Spots and Themes on Virtual Reality Technology Study in Education	2016
Incorporating Behavioral Trust Theory into System Development for Ubiquitous Applications	2014
A Model for Mining Relevant and Non-redundant Information	2012
RefMed: Relevance Feedback Retrieval System Fo PubMed	2009
Mining Team Characteristics to Predict Wikipedia Article Quality	2016
Unique Qualities of Historians' Information-seeking Behavior in Historical Research	2010
Towards a Pattern Language for Construction and Maintenance of Software Architecture Traceability Links	2016
Productive Love: A New Approach for Designing Affective Technology	2009
Recommendation Approaches for e-Learners: A Survey	2015
Quality Factory and Quality Notification Service in Data Warehouse	2010
Managing Software Quality Through a Hybrid Defect Content and Effectiveness Model	2008
Analyzing the Effectiveness of Pruning and Grouping Methods Used in Literature-Based Discovery Tools	2010
An Approach for Mining Similarity Profiled Temporal Association Patterns Using Gaussian Based Dissimilarity Measure	2015
Leveraging Knowledge Bases for Contextual Entity Exploration	2015
Computer-supported Collaborative Learning and Conceptual Change	2007
Workshop on Internet Economics (WIE2013) Report	2014
An Overview of Evolutionary Computing for Interpretation in the Oil and Gas Industry	2016
Scalable Recommender Systems: Where Machine Learning Meets Search	2015
Anticipating Annotations and Emerging Trends in Biomedical Literature	2008
A Bayesian Logistic Regression Model for Active Relevance Feedback	2008
The ATB Framework: Quantifying and Classifying Epistemic Strategies in Tangible Problem-Solving Tasks	2015
Big Data and Transparency: Using MapReduce Functions to Increase Public Expenditure Transparency	2016
Using Clinical Preferences in Argumentation About Evidence from Clinical Trials	2010
Text and Data Mining Techniques in Adverse Drug Reaction Detection	2015
Heavy Hitter Estimation over Set-Valued Data with Local Differential Privacy	2016
Dynamic Analysis of Ada Programs for Comprehension and Quality Measurement	2008
Dynamic Analysis of Ada Programs for Comprehension and Quality Measurement	2008
RecLab: A System for eCommerce Recommender Research with Real Data, Context and Feedback	2011
Design Pattern Based Decision Support	2011
LITSEEK: Public Health Literature Search by Metadata Enhancement with External Knowledge Bases	2009
Diversifying Query Suggestions Based on Query Documents	2014
A Model-based and Simulation-assisted FMEDA Approach for Safety-relevant E/E Systems	2015
Digital Wildfires: Propagation, Verification, Regulation, and Responsible Innovation	2016
TILD: A Strategy to Identify Cancer-related Genes Using Title Information in Literature Data	2014
Simple Dynamic Emission Strategies for Microblog Filtering	2016
Exploring Structural Differences in Thesauri for SKOS-based Applications	2011
Recommending New Items to Ephemeral Groups Using Contextual User Influence	2016
Opportunities for Computing to Support Healthy Sleep Behavior	2010
Integrating Variability and Safety Analysis Models Using Commercial UML-based Tools	2015
A Cross-media Model for Automatic Image Annotation	2014
Impact of Surrogate Assessments on High-Recall Retrieval	2015
Understanding Digital Badges Through Feedback, Reward, and Narrative: A Multidisciplinary Approach to Building Better Badges in Social Environments	2015
A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP	2015
A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP	2015
The Economics of Data: Quality, Value & Exchange in Web Observatories	2013
Image Tag Assignment, Refinement and Retrieval	2015
The Information Network: Exploiting Causal Dependencies in Online Information Seeking	2016
gSVD++: Supporting Implicit Feedback on Recommender Systems with Metadata Awareness	2013
Modeling and Validating Self-adaptive Service-oriented Applications	2015
Model-driven Derivation of Product Architectures	2007
Is a Picture Really Worth a Thousand Words?: - on the Role of Images in e-Commerce	2014
Exploring Smart Grid and Data Center Interactions for Electric Power Load Balancing	2014
A Hybrid Prediction Model with F-score Feature Selection for Type II Diabetes Databases	2010
Knowledge Management Instruments	2011
Efficient Interaction Between OS and Architecture in Heterogeneous Platforms	2011
Formal Specification Comprehension: The Art of Reading and Writing Z	2014
Cloud Mobile Gaming: Modeling and Measuring User Experience in Mobile Wireless Networks	2012
Characterizing the Uncertainty of Web Data: Models and Experiences	2011
The Socio-technical Design of a Library and Information Science Collaboratory	2010
A Novel Data Mining Approach for Detecting Spam Emails Using Robust Chi-Square Features	2015
Multimedia Search Reranking: A Literature Survey	2014
Rapid Requirements Checks with Requirements Smells: Two Case Studies	2014
A Web-based Photo Management System for Large Photo Collections with User-customizable Quality Assessment	2011
An Information Theoretic Approach to Sentiment Polarity Classification	2012
Recommendation in Heterogeneous Information Networks with Implicit User Feedback	2013
Techniques and Devices Used in Palpation Simulation with Haptic Feedback	2016
Regularising Factorised Models for Venue Recommendation Using Friends and Their Comments	2016
FeatureSmith: Automatically Engineering Features for Malware Detection by Mining the Security Literature	2016
Is Wikipedia Link Structure Different?	2009
Open Access Repositories As Channel of Publication Scientific Grey Literature	2015
SimPal: A Design Study on a Framework for Flexible Safety-critical Software Development	2013
Conceptualizing Documents with Wikipedia	2012
Supporting Human Memory in Personal Information Management	2008

Title	Year
Tapping into Knowledge Base for Concept Feedback: Leveraging Conceptnet to Improve Search Results for Difficult Queries	2012
Reviewing the Landscape of Research on the Threats to the Quality of User-generated Content	2016
Design with the Feet: Walking Methods and Participatory Design	2014
Toward Identification and Adoption of Best Practices in Algorithmic Recommender Systems Research	2013
Feature Location for Software Product Line Migration: A Mapping Study	2014
Enriching User Profiling with Affective Features for the Improvement of a Multimodal Recommender System	2009
Problem Characterization and Abstraction for Visual Analytics in Behavior-based Malware Pattern Analysis	2014
DoKnowMe: Towards a Domain Knowledge-driven Methodology for Performance Evaluation	2016
Performance and Dependability Modeling with M{o}Bius	2009
Revisiting the Theory of Planned Behavior for the Preparation of the Adoption of Municipal e-Services in Less Developed Countries	2012
Experiments on Density-Constrained Graph Clustering	2015
The Keepup Recommender System	2007
Semantic Tagging for Managing Web-based Learning Resources: Models, Methods and a Plattform for Supporting Resource-based Learning	2011
Locating Sensors in the Wild: Pursuit of Ranging Quality	2010
Impact of Classes of Development Coordination Tools on Software Development Performance: A Multinational Empirical Study	2008
Designing a Mobile Application to Support the Indicated Prevention and Early Intervention of Childhood Anxiety	2015
NITOS Energy Monitoring Framework: Real Time Power Monitoring in Experimental Wireless Network Deployments	2014
Adaptive Display Power Management for OLED Displays	2012
A Segmentation-based Method to Extract Structural and Evolutionary Features for Protein Fold Recognition	2014
An Empirical Study on Quality Issues of Production Big Data Platform	2015
Mining Novel Knowledge from Biomedical Literature Using Statistical Measures and Domain Knowledge	2016
Community Membership Identification from Small Seed Sets	2014
Verification of Data Pattern for Interactive Privacy Preservation Model	2011
Label-checking Strategies to Adapt Behaviour to Design	2007
Towards Analyzing Eye Tracking Data for Evaluating Interactive Visualization Systems	2014
Summarizing Biological Literature with BioSumm	2010
Leverage the Associations Between Documents, Subject Headings and Terms to Enhance Retrieval	2014
A Boosted SVM Based Sentiment Analysis Approach for Online Opinionated Text	2013
Quantitative Performance Evaluation on Segmentation Methods for SAR Ship Images	2010
Case Study on the Business Value Impact of Personalized Recommendations on a Large Online Retailer	2012
Tailoring Software Development	2010
Assessment of User Needs for Self-management Services in Coronary Heart Disease: A Designerly Approach	2010
Socializing the Semantic Gap: A Comparative Survey on Image Tag Assignment, Refinement, and Retrieval	2016
Environmental Impact of IEEE 802.11 Access Points: A Case Study	2011
Quality in Use of Domain-specific Languages: A Case Study	2011
Visual-Based Classification of Figures from Scientific Literature	2015
An Approach of Generating Test Requirements for Agile Software Development	2015
Beyond Position Bias: Examining Result Attractiveness As a Source of Presentation Bias in Clickthrough Data	2010
Benchmarking Declarative Approximate Selection Predicates	2007
HelpMeter: A Nonlinear Model for Predicting the Helpfulness of Online Reviews	2008
Text Mining for Bone Biology	2010
Beyond Usability and Performance: A Review of User Experience-focused Evaluations in Visualization	2016
Context-based Literature Digital Collection Search	2009
Advanced Visual Analytics Methods for Literature Analysis	2012
Discovering Disease-associated Drugs Using Web Crawl Data	2016
Evaluating the Evaluations of Code Recommender Systems: A Reality Check	2016
A Service Architecture for Sensor Data Provisioning for Context-aware Mobile Applications	2008
An Investigation of 'Soft' Operations Research Methods to Inform Hybrid Simulation Studies on Environmental Disasters	2015
Towards Multilingual User Models for Personalized Multilingual Information Retrieval	2011
Experience Building Non-functional Requirement Models of a Complex Industrial Architecture (Abstracts Only)	2011
Biblical Intertextuality in a Digital World: The Tool GERTRUDE	2013
Towards Context-aware Personalization and a Broad Perspective on the Semantics of News Articles	2010
Textbooks for Developing Regions	2012
Protecting Data Privacy from Being Inferred from High Dimensional Correlated Data	2014
Using Mobile Message to Improve Student Participation in Blended Courses: A Brazilian Case Study	2016
Analysis of Decision Support Systems Implemented for Suggesting a Course for Higher Education	2016
The Barriers Facing Women in the Information Technology Profession: An Exploratory Investigation of Ahuja's Model	2014
Real-use Evaluation of Effects: Emergency Departments Aiming for 'Warm Hands'	2012
Real-world Performance Modelling of Enterprise Service Oriented Architectures: Delivering Business Value with Complexity and Constraints (Abstracts Only)	2011
Towards Higher Relevance and Serendipity in Scholarly Paper Recommendation by Kazunari Sugiyama and Min-Yen Kan with Martin Vesely As Coordinator	2015
Interference Management for Distributed Parallel Applications in Consolidated Clusters	2016
Model-Driven Approach to Software Architecture Design	2009
Comparison of Collaborative Filtering Algorithms: Limitations of Current Techniques and Proposals for Scalable, High-performance Recommender Systems	2011
Have We Missed Something?: Identifying Missing Types of Research in Computing Education	2010
A Graph-based Approach to Measuring Semantic Relatedness in Ontologies	2011
On Learning Researchers' Dynamic Information Needs: An Entropy-Based Query Expansion Approach	2012
Fidelity Reloaded	2008
EXPERIENCE: Glitches in Databases, How to Ensure Data Quality by Outlier Detection Techniques	2016
PERIKLES: A Workflow Based Approach to Support the Operating Room Management	2011
How Temporal Work Styles and Product Modularity Influence Software Quality and Job Satisfaction	2016
A Survey of Context Data Distribution for Mobile Ubiquitous Systems	2012
Mining Disease-specific Molecular Association Profiles from Biomedical Literature: A Case Study	2008

Title	Year
An Integrated Data Mining Approach to Real-time Clinical Monitoring and Deterioration Warning	2012
Users Sleeping Time Analysis Based on Micro-blogging Data	2012
Content is Still King: The Effect of Neighbor Voting Schemes on Tag Relevance for Social Image Retrieval	2012
Automatically Building Research Reading Lists	2010
Hierarchical Service Analytics for Improving Productivity in an Enterprise Service Center	2010
A First Implementation and Evaluation of the IEEE 802.11Aa Group Addressed Transmission Service	2013
Identification of the Main Methods Used in Simulation Projects	2015
Issue-Focused Documentaries Versus Other Films: Rating and Type Prediction Based on User-Authored Reviews	2016
Bottom-up Adoption of Software Product Lines: A Generic and Extensible Approach	2015
An Iterated Graph Laplacian Approach for Ranking on Manifolds	2011
Enforcing K-anonymity in Web Mail Auditing	2016
Prediction of Scaffold Proteins Based on Protein Interaction and Domain Architectures	2015
My Reviewers: Participatory Design & Crowd-sourced Usability Processes	2015
Survey of Temporal Information Retrieval and Related Applications	2014
Usability Evaluation Guidelines for Business Intelligence Applications	2013
Tagging Wikipedia: Collaboratively Creating a Category System	2012
A Gender Lens Perspective of the Use of Social Network in Higher Education in Malaysia and Australia	2014
A Fast and Simple Method for Extracting Relevant Content from News Webpages	2009
California Fault Lines: Understanding the Causes and Impact of Network Failures	2010
California Fault Lines: Understanding the Causes and Impact of Network Failures	2010
Intangible Cultural Heritage Preservation: An Exploratory Study of Digitization of the Historical Literature of Chinese Kunqu Opera Librettos	2014
A Survey on Vision-based Fall Detection	2015
A Quality-scalable and Energy-efficient Approach for Spectral Analysis of Heart Rate Variability	2014
Generative Modeling of Entity Comparisons in Text	2014
Shaping Unlimited Patterns: A Vision for State-of-the-Art Visual Scalability	2014
Model-based Performance Analysis of Software Architectures Under Uncertainty	2013
Fair Pattern Discovery	2014
Single Toolset Modeling Approaches in Semiconductor Manufacturing	2012
Two-tier Similarity Model for Story Link Detection	2010
Influence Maximization in Near-Linear Time: A Martingale Approach	2015
A Sensemaking Environment for Literature Study	2012
Bayesian Methods for Multimedia Signal Processing	2007
MKEM: A Multi-level Knowledge Emergence Model for Mining Undiscovered Public Knowledge	2009
How Do Students Learn to Program in a Connected World?	2013
Scientific Articles Recommendation	2013
Opportunities for Computing Technologies to Support Healthy Sleep Behaviors	2011
Infrastructuring in Participatory Design	2014
Unveiling Anomalies and Their Impact on Software Quality in Model-based Automotive Software Revisions with Software Metrics and Domain Experts	2016
Unsupervised Multi-feature Tag Relevance Learning for Social Image Retrieval	2010
Rethinking User Generated Location Rating: Where Does the Lion Get Its Share?	2015
A Taxonomy for Requirements Engineering and Software Test Alignment	2014
Descriptive Metadata, Iconclass, and Digitized Emblem Literature	2012
Efficient Algorithms for Genome-wide Association Study	2009
Whetting the Appetite of Scientists: Producing Summaries Tailored to the Citation Context	2009
Integrating Decision Making Methods on Requerimentos Elicitation Process	2016
Prioritizing Adverse Drug Reaction and Drug Repositioning Candidates Generated by Literature-Based Discovery	2016
Complex Network Perspective on Network Dynamics: A Study on Investment Network in VC Industry in China	2012
GLMix: Generalized Linear Mixed Models For Large-Scale Response Prediction	2016
Tailored Scenarios: A Low-cost Online Method to Elicit Perceptions on Designs Using Real Relationships	2013
Mining Research Abstracts for Exploration of Research Communities	2012
Combining Syntactic Information and Domain-specific Lexical Patterns to Extract Drug-drug Interactions from Biomedical Texts	2010
Interactive Wifi Connectivity for Moving Vehicles	2008
Interactive Wifi Connectivity for Moving Vehicles	2008
Soft or Approaches in Problem Formulation Stage of a Hybrid M&S Study	2014
QoE Doctor: Diagnosing Mobile App QoE with Automated UI Control and Cross-layer Analysis	2014
Quality Software Development: What Do We Need to Improve in the Software Development Process?	2008
Detecting Conversing Groups with a Single Worn Accelerometer	2014
Product Line Bundles to Support Product Derivation in Multi Product Lines	2011
k-Information Gain Scaled Nearest Neighbors: A Novel Approach to Classifying Protein-Protein Interaction-Related Documents	2012
A Visual Analytics Approach to Finding Factors Improving Automatic Speaker Identifications	2015
A Preprocessing Method of Internet Search Data for Prediction Improvement: Application to Chinese Stock Market	2012
Hybrid Framework Using Multiple-filters and an Embedded Approach for an Efficient Selection and Classification of Microarray Data	2016
A Framework for Source Code Metrics	2010
What's in This Paper?: Combining Rhetorical Entities with Linked Open Data for Semantic Literature Querying	2015
A Survey of Measures and Methods for Matching Geospatial Vector Datasets	2016
Survey on Web Spam Detection: Principles and Algorithms	2012
An Industrial Case Study on Quality Impact Prediction for Evolving Service-oriented Software	2011
Facilitating Participatory Practice for Smart Textiles	2015
Leakage in Data Mining: Formulation, Detection, and Avoidance	2012
Predicting Information Technology Usage in the Car: Towards a Car Technology Acceptance Model	2012
A Comparative Study of Different Heavy Tail Index Estimators of the Flow Size from Sampled Data	2007
Linking Questions and Evidence	2008
The Need for a Multi-factorial Model of Safe Human Performance in Air Traffic Control	2010

Title	Year
E-governance and Governance: A Case Study of the Assessment of the Effects of Integrated Financial Management System on Good Governance in Two Municipal Councils in Kenya	2007
Reconsideration of the Simulated Work Task Situation: A Context Instrument for Evaluation of Information Retrieval Interaction	2010
Towards a Privacy Threat Model for Public Displays	2015
A Measurement Study Regarding Quality of Service and Its Impact on Multiplayer Online Games	2010
DyPS: Dynamic Processor Switching for Energy-aware Video Decoding on Multi-core SoCs	2014
An Exploration of Tool Support for Categorical Coding	2008
Evaluating Teamwork Support in Tabletop Groupware Applications Using Collaboration Usability Analysis	2008
Reducing Long Queries Using Query Quality Predictors	2009
Adapting Collaborative Radiological Practice to Low-resource Environments	2012
High-Density Image Storage Using Approximate Memory Cells	2016
Thinking in Stories: Narrative-based Design for Higher-Level Thinking	2016
Associating Architectural Issues with Quality Attributes: A Survey on Expert Agreement	2016
Guidelines for Multilingual Linked Data	2013
A Social Model for Literature Access: Towards a Weighted Social Network of Authors	2010
A Method Based on Naming Similarity to Identify Reuse Opportunities	2016
Multi-agent Simulation of Ground Aircraft Operations at a Large Airport	2010
A Framework for Supporting User-centric Collaborative Information Seeking	2011
Time-series Data Mining	2012
Using Canonical Correlation Analysis for Generalized Sentiment Analysis, Product Recommendation and Search	2011
User-interface Design for Highly Automated Systems a Structured Approach	2013
Probabilistic Models for Topic Learning from Images and Captions in Online Biomedical Literatures	2009
Optimizing Cost and Quality by Integrating Inspection and Test Processes	2011
Moving Beyond Sharing vs. Withholding to Understand How Scientists Share Data Through Large-scale, Open Access Databases	2011
Profiling Relational Data: A Survey	2015
Monitoring Code Quality and Development Activity by Software Maps	2011
Work and Play: An Experiment in Enterprise Gamification	2016
A Study on Service Identification Methods for Software Product Lines	2012
A Survey of View Selection Methods	2012
Geospatial Data Qualities As Web Services Performance Metrics	2007
A Survey of Software Aging and Rejuvenation Studies	2014
Dynamic Reconfiguration in Multi-hop Switched Ethernet Networks	2014
Survey: Gap-filling Methods for the Positron Emission Tomograph Sinograms	2011
When is Collaborating with Friends a Good Idea? Insights from Design Education	2009
Citation Chain Aggregation: An Interaction Model to Support Citation Cycling	2011
Investigating Code Reading Techniques for Novice Inspectors: An Industrial Case Study	2014
Demonstrating Social Support from Autism Bloggers Community on Twitter	2015
Assessing Data Mining Results via Swap Randomization	2007
Fastest Association Rule Mining Algorithm Predictor (FARM-AP)	2011
Ludic Aspects of the Generative Audiovisual Narrative System	2008
Optimally Discriminative Choice Sets in Discrete Choice Models: Application to Data-Driven Test Design	2016
Interacting with In-vehicle Systems: Understanding, Measuring, and Evaluating Attention	2009
QDSL: A Queuing Model for Systems with Differential Service Levels	2008
Sizing Sketches: A Rank-based Analysis for Similarity Search	2007
FISH: A Practical System for Fast Interactive Image Search in Huge Databases	2008
Task Migration for Fault-tolerance in Mixed-criticality Embedded Systems	2009
Empirical Research on the Total Factor Productivity of Chinese Software Companies	2012
When the Levee Breaks: Without Bots, What Happens to Wikipedia's Quality Control Processes?	2013
Charting the Digital Library Evaluation Domain with a Semantically Enhanced Mining Methodology	2013
Providing Statistically Guaranteed Streaming Quality for Peer-to-peer Live Streaming	2009
'My World(s)': A Tabletop Environment to Support Fantasy Play for Kindergarten Children	2007
Towards a Better Understanding of Simplicity in Agile Software Development Projects	2016
Towards a Quality Model for Model Composition Effort	2014
Two-phase Prediction of Protein Functions from Biological Literature Based on Gini-Index	2011
Maximizing Acceptance Probability for Active Friending in Online Social Networks	2013
Curiosity: From Psychology to Computation	2013
Blind Evaluation: Student's Experience of the Empathic Lead User Method	2007
Towards Mutation Analysis for Use Cases	2016
Optimization of Cross-docking Terminal Using Flexsim/Optquest: Case Study	2014
Scalable Community Discovery on Textual Data with Relations	2008
Leakage in Data Mining: Formulation, Detection, and Avoidance	2011
'Desperately' Seeking Research on Agile Requirements in the Context of Large-scale Agile Projects	2015
Agents Assisted Software Project Management	2008
Collecting Legacy Corpora from Social Science Research for Text Mining Evaluation	2010
Twitter Analytics: A Big Data Management Perspective	2014
A Data Mining Approach to MPGN Type II Renal Survival Analysis	2010
Exploring Language Support for Immutability	2016
From Machu_Picchu to Rafting the Uru'bamba River: Anticipating Information Needs via the Entity-query Graph	2013
Supporting Mobile Sensemaking Through Intentionally Uncertain Highlighting	2016
On Numerical Modeling Performance of Generalized Preconditioned Methods	2013
Monotonicity and Stratification	2008
Multiple Feature Fusion for Social Media Applications	2010
Size Estimation of Non-cooperative Data Collections	2012
Evaluating Joint Modeling of Yeast Biology Literature and Protein-protein Interaction Networks	2012
The Past, Present, and Future of MOOCs and Their Relevance to Software Engineering	2014

Title	Year
Social Reference: Aggregating Online Usage of Scientific Literature in CiteULike for Clustering Academic Resources	2011
An Event-based Framework for Characterizing the Evolutionary Behavior of Interaction Graphs	2009
The Evaluation of a Constructive Modeling and Simulation Approach in Teaching Port Management Skills	2015
Discovering Discriminative Graph Patterns from Gene Expression Data	2016
Assessing Topic Discovery Evaluation Measures on Facebook Publications of Political Activists in Brazil	2016
Views on Internal and External Validity in Empirical Software Engineering	2015
Automated Generation of Music Playlists: Survey and Experiments	2014
Supporting Pilgrims in Navigating Densely Crowded Religious Sites	2012
Tactics for Homing in Mobile Life: A Fieldwalk Study of Extremely Mobile People	2010
Arabic Text Categorization Based on Arabic Wikipedia	2014
A Framework for Characterization and Analysis of Software System Scalability	2007
Monitoring Algorithms for Negative Feedback Systems	2010
An Approach to Customizing Requirements Goal Model Based on Metamodel for Ontology Registration	2009
Designing Interactive Systems for the Experience of Time	2013
Looking Back on the Current Day: Interruptibility Prediction Using Daily Behavioral Features	2016
Wrangler: Interactive Visual Specification of Data Transformation Scripts	2011
An Approach to Join Ontologies and Their Reuse in the Construction of Application Ontologies	2014
Non-intrusive Detection of Psycho-social Dimensions Using Sociolinguistics	2013
A Framework to Integrate ICT Security Awareness into the South African Schooling System	2014
Verifying SystemC: A Software Model Checking Approach	2010
A Linear-time Approach for the Transient Thermal Simulation of Liquid-cooled 3D Ics	2011
Construct Anticancer Drug-drug Correlation Network	2009
Level Up: A Frame Work for the Design and Evaluation of Educational Games	2009
Towards Biophysical Validation of Constraint Modeling for Rigidity Analysis of Proteins	2012
Towards a Framework for the Adoption of Business Intelligence in Public Sector Organisations: The Case of South Africa	2011
Dependencies Between E-Learning Usage Patterns and Learning Results	2013
Ad Click Prediction: A View from the Trenches	2013
A Semantic Oriented Method for Conceptual Data Modeling in ontoUML Based on Linguistic Concepts	2011
Measuring and Fingerprinting Click-spam in Ad Networks	2012
MedRank: Discovering Influential Medical Treatments from Literature by Information Network Analysis	2013
On Clinical Decision Support	2014
A New Data Structure for Asynchronous Periodic Pattern Mining	2009
Automatic Assessment of Website Compliance to the European Cookie Law with CooLCheck	2016
An Extended Case Study Methodology for Investigating Influence of Cultural, Organizational, and Automation Factors on Human-automation Trust	2014
An Agent-based Approach to Modeling Online Social Influence	2013
Localization of Wireless Sensor Networks in the Wild: Pursuit of Ranging Quality	2013
Formalizing Correspondence Rules for Automotive Architecture Views	2014
Mitigating Slow Dynamics of Low-Cost Chemical Sensors for Mobile Air Quality Monitoring Sensor Networks	2016
Power and Bias of Subset Pooling Strategies	2007
Discovering Organizational Perspective in Workflow Using Agent Approach: An Illustrative Case Study	2010
Peer Production of Structured Knowledge -: An Empirical Study of Ratings and Incentive Mechanisms	2008
The Impact of Test Ownership and Team Structure on the Reliability and Effectiveness of Quality Test Runs	2014
Brief Author Index	2009
Grid-based Hierarchy Structure for Mining and Querying Vehicular Ad-hoc Networks	2012
ADSEng: A Model-based Methodology for Autonomous Digital Service Engineering	2016
Comparing Semantic Associations in Sentences and Paragraphs for Opinion Detection in Blogs	2009
Collaboration, Computation, and Creativity: Media Arts Practices in Urban Youth Culture	2007
Clustering Memes in Social Media	2013
Achieving Scalable Model-based Testing Through Test Case Diversity	2013
Simple Time-biased KNN-based Recommendations	2010
Achieving Long-term Operation with a Capacitor-driven Energy Storage and Sharing Network	2012
Semantic Representation of Context Models: A Framework for Analyzing and Understanding	2009
People Like Us: Mining Scholarly Data for Comparable Researchers	2014
Cloud Services Evaluation Framework	2012
Teaching Programming with Flipped Classroom Method: A Study from Two Programming Courses	2015
Analysis of Tactile Elements Based on Classification of a Tactile User Interface	2016
Relevance Scores for Triples from Type-Like Relations	2015
Achieving ICS Resilience and Security Through Granular Data Flow Management	2016
An Overview of Context Types Within Multimedia and Social Computing	2015
Theoretical Foundations to Information Technology Outsourcing Research in the Public Sector	2016
Identifying Software Process Management Challenges: Survey of Practitioners in a Large Global IT Company	2015
Multinational Perspectives on Information Technology from Academia and Industry	2015
IFME: Information Filtering by Multiple Examples with Under-sampling in a Digital Library Environment	2013
Research Issues in Software Fault Categorization	2007
Spatial Diversity, Do Users Appreciate It?	2010
Reducing the Effort of Bug Report Triage: Recommenders for Development-oriented Decisions	2011
Detection and Classification of Peer-to-peer Traffic: A Survey	2013
Knowledge Representation in ICU Communication	2010
A Report on a Survey and Study of Static Analysis Users	2008
Making Interactive Systems More Flexible: An Approach Based on Users' Participation and Norms	2010
Like It or Not: A Survey of Twitter Sentiment Analysis Methods	2016
Unbiased Ranking Evaluation on a Budget	2015
Educational Data Mining Methods for Studying Student Behaviors Minute by Minute Across an Entire School Year	2010
Fast Algorithms for Approximating the Singular Value Decomposition	2011
An Empirical Study on Investigating the Role of KMS in Promoting Trust Within GSD Teams	2013
Axiomatic Analysis and Optimization of Information Retrieval Models	2013

Title	Year
Formation of Voluntary Inter-organizational IT Governance for Healthcare and Social Welfare IT: Theoretical Background and Empirical Evaluation	2014
Know Your Enemy: Compromising Adversaries in Protocol Analysis	2014
Sensitivity of Results to Different Data Quality Meta-data Criteria in the Sample Selection of Projects from the ISBSG Dataset	2010
Cloud-based Support for Global Software Engineering: Potentials, Risks, and Gaps	2014
Verification and Validation of Interactive Systems in the CMMI-DEV Capability Maturity Model	2016
Synthesis of Secure Software Development Controls	2015
Towards a Model for Cost-benefit-analysis of Quality Assurance in the Automotive E/E Development	2008
Architectural Stability Evolution in Open-Source Systems	2015
Best Practices for Capturing Context in User Experience Studies in the Wild	2011
More Bang for Their Bucks: Assessing New Features for Online Advertisers	2007
A Proactive Recommendation System for Writing: Helping Without Disrupting	2007
Towards Usable Cyber Security Requirements	2009
Active Objects: Actions for Entity-centric Search	2012
Supporting Children with Special Needs Through Multi-perspective Behavior Analysis	2011
Sub-document Level Information Retrieval: Retrieval and Evaluation	2012
Looking but Not Seeing: Implications for HCI	2007
Supporting Oracle Construction via Static Analysis	2016
Potential Individual Differences Regarding Automation Effects in Automated Driving	2014
Optimizing Memory Bandwidth Use and Performance for Matrix-vector Multiplication in Iterative Methods	2011
Surveys in Software Engineering: Identifying Representative Samples	2016
A Survey on Ear Biometrics	2013
Decision Making in Autonomic Computing Systems: Comparison of Approaches and Techniques	2011
Semantic Text Mining for Lignocellulose Research	2011
Digital Democracy: An Analysis of Resources and Acceptance	2016
Emergent Participant Interaction	2012
Electronic Portfolios As Tools for Organizing Student Information	2008
Method and Developer Characteristics for Effective Agile Method Tailoring: A Study of XP Expert Opinion	2010
Dynamic Design Elements for the Peripheral Interaction of Ambient Media	2008
Bridge the Gap!: What Can Work Design in Crowdwork Learn from Work Design Theories?	2015
Beyond Data: Contextual Information Fusion for Cyber Security Analytics	2016
Weighted Graph Comparison Techniques for Brain Connectivity Analysis	2013
Looking at Near-duplicate Videos from a Human-centric Perspective	2010
An Approach for Personalized Web Information Retrieval Using Modified PageRank Method	2016
Patient Health Portal: A Calendar Paradigm	2012
Estimating Entity Importance via Counting Set Covers	2012
Exploring the Design-Space: The Authorial Game Evolution Tool Case-Study	2016
Incremental Tensor Analysis: Theory and Applications	2008
Consumer Decision Patterns Through Eye Gaze Analysis	2010
Software and Systems Engineering Process Implications of Evolutionary Development (Keynote Abstract)	2011
Evaluation of Approaches for Designing Secure Data Warehouse	2012
Using Semi-parametric Clustering Applied to Electronic Health Record Time Series Data	2011
Introducing a Text Annotation Tool (OntoMate), Assisting Curation at Rat Genome Database	2016
Collective Spatial Keyword Queries: A Distance Owner-driven Approach	2013
Enriching Living Lab-approaches for ICT-innovation by Introducing Different User Roles: The Case of Digital TV	2010
A Survey of Handoff Schemes for Vehicular Ad-hoc Networks	2010
Designing Exploratory Search Tasks for User Studies of Information Seeking Support Systems	2009
When Food Quality Control in China Meets Mobile and Wireless Technology: Interactions, Potentials and Pitfalls	2013
Machine Learning and Causal Inference for Policy Evaluation	2015
Supporting the Supermarket Shopping Experience Through a Context-aware Shopping Trolley	2009
Mobile Phone User Interface Design for Patients with Traumatic Brain Injury	2010
Enriching How-to Guides by Linking Actionable Phrases	2016
Empirical Comparison of Four Classifier Fusion Strategies for Positive-versus-negative Ensembles	2011
Students' Use of Mobile Technology to Collect Data in Guided Inquiry on Field Trips	2013
Engineering Web Information Systems: A Content Management System-based Approach	2011
Reliability and Survivability Methodologies for Next Generation Networks	2008
How Do Public Library Administrators Generate and Evaluate Ideas for New Services?: A Proposed Model Based on Evidence from Cuyahoga County Public Library	2010
Data-driven Compound Splitting Method for English Compounds in Domain Names	2009
Correlation of Node Importance Measures: An Empirical Study Through Graph Robustness	2015
Usability Dimensions for Content-Orientation Techniques in Multi-Touch Tabletop Displays: An Overview	2016
Preliminary Design Examination of the ParalleX System from a Software and Hardware Perspective	2011
Choices and Constraints: Research Goals and Approaches in Information Retrieval (Part 2)	2014
Mining Opinions in Comparative Sentences	2008
A Survey on Smartphone-based Systems for Opportunistic User Context Recognition	2013
An Approach for Implementing Core Assets in Service-oriented Product Lines	2011
An Axiomatic Approach to Personalized Ranking Systems	2010
Fastanova: An Efficient Algorithm for Genome-wide Association Study	2008
Simulations in Wireless Sensor and Ad Hoc Networks	2007
A Qualitative Metasynthesis of Activity Theory in SIGDOC Proceedings 2001-2011	2012
A Data-driven Approach to Model Culture-specific Communication Management Styles for Virtual Agents	2010
Choices and Constraints: Research Goals and Approaches in Information Retrieval (Part 1)	2014
Studying Shared Situation Awareness by Agent-Based Simulation	2013
Friend Recommendation for Weight Loss App	2014
Designed and Delivered Today, Eroded Tomorrow?: Towards an Open and Lean Architecting Framework Balancing Agility and Sustainability	2016
ANDES: Efficient Evaluation of NOT-twig Queries in Relational Databases	2012

Title	Year
Modeling Resource Sharing for a Road-side Access Point Supporting Drive-thru Internet	2009
Towards the Design of Intelligible Object-based Applications for the Web of Things	2011
Security in Web 2.0 Application Development	2008
Similarity Matching for Uncertain Time Series: Analytical and Experimental Comparison	2011
Temporal Formation and Evolution of Online Communities	2016
Mining Latent Entity Structures from Massive Unstructured and Interconnected Data	2014
Computer Vision Based Attentiveness Detection Methods in E-Learning	2014
A Product Line Modeling and Configuration Methodology to Support Model-based Testing: An Industrial Case Study	2012
Towards Multi-modal Anticipatory Monitoring of Depressive States Through the Analysis of Human-smartphone Interaction	2016
Moses – Method for Selecting Senior Mobile Phones: Supporting Design & Choice for the Elderly	2012
A Comprehensive Survey of System Dependability for Real Time Embedded Software	2015
Computing Bayes-Nash Equilibria Through Support Enumeration Methods in Bayesian Two-Player Strategic-Form Games	2009
An Urn Occupancy Approach for Modeling the Energy Consumption of Distributed Beaconing	2014
Reuse: Reducing Test Effort	2013
A Requirements Analysis for Videogame Design Support Tools	2009
On Deciding Between Conservative and Optimistic Approaches on Massively Parallel Platforms	2010
Deep, Efficient, and Dialogic Study of Undergraduate Information Seeking and Use: A Methodological Exploration	2010
Information Behavior in Palliative Care: An Exploratory Reflection on the Diversity and Multiplicity of Contexts	2010
Hiding Co-occurring Frequent Itemsets	2009
A Qualitative Study of the Determinants of Self-managing Team Effectiveness in a Scrum Team	2011
Quantitative Analysis of Faults and Failures with Multiple Releases of Softpm	2008
Automated Identification of Synonyms in Biomedical Acronym Sense Inventories	2010
Dynamic Indexability and Lower Bounds for Dynamic One-dimensional Range Query Indexes	2009
An NFC-based O2O Service Model in Exhibition-space	2016
Relational Operators for Prioritizing Candidate Biomarkers in High-throughput Differential Expression Data	2010
Parameterized Tiling Revisited	2010
Building Text-mining Framework for Gene-Phenotype Relation Extraction Using Deep Leaning	2015
Evidence of the Presence of Bias in Subjective Metrics: Analysis Within a Family of Experiments	2014
A Large-scale Survey on the Effects of Selected Development Practices on Software Correctness	2015
Domain Matters: Bringing Further Evidence of the Relationships Among Anti-patterns, Application Domains, and Quality-related Metrics in Java Mobile Apps	2014
Fuel-cache Site-selection for Polar Research: A Summary of Results	2009
On the Importance of Venue-Dependent Features for Learning to Rank Contextual Suggestions	2014
Assessing Usability Testing for People Living with Dementia	2016
Creativity Techniques in Game Design	2007
A Taxonomy for Software Change Impact Analysis	2011
The Innovative Behaviour of Software Engineers: Findings from a Pilot Case Study	2016
Scientific Monitoring by Mining Scientific Papers	2014
Internalization, Qualitative Methods, and Evaluation	2008
A Comparative Study of Information Granulation-based Fuzzy Inference Systems Developed by Means of Space Optimization Algorithm (SOA) and Genetic Algorithms	2011
Collaborative Information Seeking: Understanding Users, Systems, and Content	2012
Towards the Definition of a Framework for the Management of Interactive Collaborative Learning Applications for Preschoolers	2015
Exploring Skin Conductance Synchronisation in Everyday Interactions	2014
Automatic Mining of Specifications from Invocation Traces and Method Invariants	2014
A Field Study Comparing Approaches to Collecting Annotated Activity Data in Real-world Settings	2015
Non-functional Requirements Analysis Modeling for Software Product Lines	2009
An Online Academic Journal System for CINDI	2011
A Coclustering Approach for Mining Large Protein-Protein Interaction Networks	2012
Bid Generation for Advanced Match in Sponsored Search	2011
Instructor-centric Source Code Plagiarism Detection and Plagiarism Corpus	2012
Collaborative Creation of Digital Tactile Graphics	2015
Culturally Adaptive Mobile Agent Dialogue to Communicate with People in Crisis Recovery	2010
Estimating the Accuracy of User Surveys for Assessing the Impact of HPC Systems	2016
A Dynamic Trust Network Based Simulation Framework for Reputation-based Service Selection	2009
Learning to Selectively Rank Patients' Medical History	2013
Journal Ranking Based on Social Information	2009
Recommending Citations for Academic Papers	2007
From Technology Engineering to Social Engineering: 15 Years of Research on Virtual Worlds	2007
Similarity Measures for Comparing Biclusterings	2014
Learning Using Anti-training with Sacrificial Data	2016
Complex Systems, Cooperative Work, and Usability	2015
Empirical Research Methods for Software Engineering	2007
Within-Network Classification Using Radius-Constrained Neighborhood Patterns	2014
Topology-Related Metrics and Applications for the Design and Operation of Wireless Sensor Networks	2014
The Effect of Assessor Error on IR System Evaluation	2010
Top-k Typicality Queries and Efficient Query Answering Methods on Large Databases	2009
A Case Study Evaluation of Maintainability and Performance of Persistency Techniques	2008
IT Governance to Fit Your Context: Two U.S. Case Studies	2010
User Experience of Academic Staff in the Use of a Learning Management System Tool	2016
A Decision Tree Approach to Data Classification Using Signal Temporal Logic	2016
A Novel Approach for Mining Similarity Profiled Temporal Association Patterns Using Venn Diagrams	2015
Senior Wellness: Practices of Community Senior Centers	2011
UAV Video Coverage Quality Maps and Prioritized Indexing for Wilderness Search and Rescue	2010
Design and Analysis of Classifier Learning Experiments in Bioinformatics: Survey and Case Studies	2012
Asynchronous Object Storage with QoS for Scientific and Commercial Big Data	2013

Title	Year
FrauDetector: A Graph-Mining-based Framework for Fraudulent Phone Call Detection	2015
A Case Study in Test Management	2007
Mobile Applications for Open Display Networks: Common Design Considerations	2013
Adaptivity Metric and Performance for Restart Strategies in Web Services Reliable Messaging	2008
An Approach to Controlling User Models and Personalization Effects in Recommender Systems	2013
Simulating the Micro-level Behavior of Emergency Department for Macro-level Features Prediction	2015
Find It if You Can: Usability Case Study of Search Engines for Young Users	2013
Adtributor: Revenue Debugging in Advertising Systems	2014
Identifying the Semantic Orientation of Foreign Words	2011
Cross-project Defect Prediction: A Large Scale Experiment on Data vs. Domain vs. Process	2009
Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search	2009
Evolving a Software Products Line for E-commerce Systems: A Case Study	2015
Social Scientists, Documents and Cyberinfrastructure: The Cobbler's Children or the Missing Masses?	2011
Investigating and Reflecting on the Integration of Automatic Data Analysis and Visualization in Knowledge Discovery	2010
Crowdsourced Semantic Annotation of Scientific Publications and Tabular Data in PDF	2015
Assessing the Usefulness of a Requirements Monitoring Tool: A Study Involving Industrial Software Engineers	2016
How Do Software Ecosystems Evolve? A Quantitative Assessment of the R Ecosystem.	2016
Designing for Video: Investigating the Contextual Cues Within Viewing Situations	2014
Design Patterns As First-class Connectors	2013
Community Mining on Dynamic Weighted Directed Graphs	2009
A Comparative Study of Various Supervised Feature Selection Methods for Spam Classification	2016
Reflections on Cultural Heritage and Digital Humanities: Modelling in Practice and Theory	2014
Towards a Diversity-minded Wikipedia	2011
Can'T See the Forest for the Trees?: A Citation Recommendation System	2013
How Do Developers React to API Deprecation?: The Case of a Smalltalk Ecosystem	2012
Evaluating the Impact of Alternative Dependency Graph Encodings on Solving Event Extraction Tasks	2010
ClimbAX: Skill Assessment for Climbing Enthusiasts	2013
Evaluation of Methods for Relative Comparison of Retrieval Systems Based on Clickthroughs	2009
Determinants of Patients' Intention to Adopt Diabetes Self-Management Applications	2015
An Exploration of Automated Grading of Complex Assignments	2016
Investigating Homophily in Online Social Networks	2010
Metro Maps of Science	2012
A Semi-supervised Hybrid System to Enhance the Recommendation of Channels in Terms of Campaign Roi	2011
Combining Information Extraction and Text Mining for Cancer Biomarker Detection	2013
Alternative Methods Using Similarities in Software Effort Estimation	2012
From Footprint to Evidence: An Exploratory Study of Mining Social Data for Credit Scoring	2016
Strategy-Based Instruction: Lessons Learned in Teaching the Effective and Efficient Use of Computer Applications	2008
Usable Object Management Approaches for Online Social Networks	2013
Understanding the Business Value of Intelligent Products for Product Development in Manufacturing Industries	2016
Information Retrieval for e-Discovery	2010
Assessing Benefit from Feature Feedback in Active Learning for Text Classification	2011
DOTAGWA: A Case Study in Web-based Architectures for Connecting Surface Water Models to Spatially Enabled Web Applications	2007
Knowledge Transformation for Cross-domain Sentiment Classification	2009
e-Government Procurement Implementation in India: Two Comparative Case Studies from the Field	2014
Structural Vulnerability Analysis of Overlapping Communities in Complex Networks	2014
Tracing SPLs Precisely and Efficiently	2012
EmoSnaps: A Mobile Application for Emotion Recall from Facial Expressions	2015
Evaluation of a Game Used to Teach Usability to Undergraduate Students in Computer Science	2015
Mining Correlated Bursty Topic Patterns from Coordinated Text Streams	2007
Cowboys, Ankle Sprains, and Keepers of Quality: How is Video Game Development Different from Software Development?	2014
Detecting Research Topics via the Correlation Between Graphs and Texts	2007
Management Anti-patterns in IT Innovation Projects	2016
Impacts of Globalization on Indian Industry: Case of Financialization in It and Non It Sectors	2014
Mining Significant Graph Patterns by Leap Search	2008
Conditional Heavy Hitters: Detecting Interesting Correlations in Data Streams	2015
Different Aspects of Trust in Ubiquitous Intelligent Transportation Systems	2010
The Relationship Between Cognitive Abilities, Well-being and Use of New Technologies in Older People	2010
Research on Government Chief Information Officer System in China: A Critical Assessment	2010
Security Assessment of Clickjacking Risks in Web Applications: Metrics Based Approach	2015
Practical and Efficient Cryptographic Enforcement of Interval-based Access Control Policies	2011
An Exploratory Study on Role-based Collaborative Business Process Modeling Approaches	2015
Online Allocation of Display Ads with Smooth Delivery	2012
Preliminary Empirical Evidence on SPrL Variability Management with EPF and SMartySPEM	2016
Identifying Potential Bioactive Compounds of Natural Products by Combining ADMET Prediction Methods	2015
Requirements Engineering in Feature Oriented Software Product Lines: An Initial Analytical Study	2012
Risk Sensitivity of Price of Anarchy Under Uncertainty	2016
An Experimental Study on Implicit Social Recommendation	2013
HMM-based Speech Segmentation: Improvements of Fully Automatic Approaches	2016
Next Step in Electronic Brainstorming: Collaborative Creativity with the Web	2011
Exploring Dynamics and Semantics of User Interests for User Modeling on Twitter for Link Recommendations	2016
Candidate Sets for Alternative Routes in Road Networks	2015
Network Under Joint Node and Link Attacks: Vulnerability Assessment Methods and Analysis	2015
Predicting the Popularity of Trending Arabic News on Twitter	2014
Using Ontologies to Discover Fact IDs	2010
Performance Anomaly Detection and Bottleneck Identification	2015
Perceptual Depth to Detect Changes That We Don'T See	2007

Title	Year
CAMY: Applying a Pet Dog Analogy to Everyday Ubicomp Products	2014
Identification Et Prise En Compte De Propri{E}T{E}s Ergonomiques Pour La Mod{E}Lisation Et La Conception De SIM	2008
A Position-Based Method for the Extraction of Financial Information in PDF Documents	2016
A Computational Framework for Finding Interestingness Hotspots in Large Spatio-temporal Grids	2014
An Introduction to Metabolic Networks and Their Structural Analysis	2008
Statistical and Network Analysis of Metabolomics Data	2016
Supporting Developers' Coordination in the IDE	2015
The Truth, The Whole Truth, and Nothing But the Truth: A Pragmatic Guide to Assessing Empirical Evaluations	2016
Amalgamating Source Code Authors, Maintainers, and Change Proneness to Triage Change Requests	2014
Social Scientists and Cyberinfrastructure: Insights from a Document Perspective	2012
Compiler-Driven Software Speculation for Thread-Level Parallelism	2015
A Generalized Co-HITS Algorithm and Its Application to Bipartite Graphs	2009
Using Trace Sampling Techniques to Identify Dynamic Clusters of Classes	2007
A Novel Click Model and Its Applications to Online Advertising	2010
What Does Touch Tell Us About Emotions in Touchscreen-Based Gameplay?	2012
Team-based Learning in a Data Structures Course	2014
Categorization of Real-time Software Components for Code Size Estimation	2010
A Prediction-based Detection Algorithm Against Distributed Denial-of-service Attacks	2009
School of Phish: A Real-world Evaluation of Anti-phishing Training	2009
A Study of Ripple Effects in Software Ecosystems (NIER Track)	2011
Analysis of Long Queries in a Large Scale Search Log	2009
Building an Open Data Ecosystem: An Irish Experience	2014
Analysis of Schemas with Access Restrictions	2015
Beyond Linear Chain: A Journey Through Conditional Random Fields for Information Extraction from Text	2014
Biomedical Association Mining and Validation	2010
Community Detection from Location-tagged Networks	2014
Information Practices of Disaster Preparedness Professionals in Multidisciplinary Groups	2010
A Comparative Study of DASH Representation Sets Using Real User Characteristics	2016
Scalability Studies on Selective Mutation Testing	2015
TCSST: Transfer Classification of Short & Sparse Text Using External Data	2012
Artifacts of Software Reference Architectures: A Case Study	2014
Measuring Situation Awareness in Command and Control: Comparison of Methods Study	2007
A Case Study of Applying Software Product Line Engineering to the Air Conditioner Domain	2016
Integrating Document Clustering and Multidocument Summarization	2011
Is It All About Awareness?: People, Smart Cities 3.0, and Evolving Spaces for IT	2016
Fast Recovery from Dual-link or Single-node Failures in IP Networks Using Tunneling	2010
Oasis: A Framework for Linking Notification Delivery to the Perceptual Structure of Goal-directed Tasks	2010
Exact and Inexact Methods for Selecting Views and Indexes for OLAP Performance Improvement	2008
Exploring Underproduction in Wikipedia	2011
Telecom Industry in the Cloud Value Chain: Methods for Assessing Business Opportunities	2012
Learning Transformations for Clustering and Classification	2015
Making the Production of Learning at Scale More Open and Flexible	2016
Designing Auditory Cues to Enhance Spoken Mathematics for Visually Impaired Users	2010
Using Domain Ontologies for Finding Experts in Corporate Wikis	2011
MOE Quantification of Missions Using Sensor Data Driven Graph Similarity Metrics	2015
A Conceptual Model of Client-driven Agile Requirements Prioritization: Results of a Case Study	2010
Automatic Detection of Cyberbullying on Social Networks Based on Bullying Features	2016
Evaluation of Multimedia Services in Mobile WiMAX	2008
Investigating the Use of Grounded Theory in Information Systems Research	2008
Tag Recommendation for Georeferenced Photos	2011
Size and Cohesion Metrics As Indicators of the Long Method Bad Smell: An Empirical Study	2015
Metrics-driven Security Objective Decomposition for an e-Health Application with Adaptive Security Management	2013
Does Service Composition Suffice to Define Business Contracts for IT Services in Networked Organizations	2013
Learning to Rank Results in Relational Keyword Search	2011
Challenges in Computing Semantic Relatedness for Large Semantic Graphs	2014
Modeling Twitter Acceptance and Use Under the Risk of Antisocial Behavior	2016
An Empirical Analysis of Semantic Network in Online Crowdfunding: Evidence from Kickstarter	2016
Progressive Light Transport Simulation on the GPU: Survey and Improvements	2014
ParaXplore Interfaces: Parametric Interfaces for Parallel Exploration in Design	2016
TiNYARM: Awareness of Research Papers in a Community of Practice	2013
Methodological Approaches to the Evaluation of Game Music Systems	2014
Towards GeoSpatial Semantic Data Management: Strengths, Weaknesses, and Challenges Ahead	2014
Advertiser-centric Approach to Understand User Click Behavior in Sponsored Search	2011
How to Evaluate Technologies for Health Behavior Change in HCI Research	2011
Taxonomy-superimposed Graph Mining	2008
Generative Data Models for Validation and Evaluation of Visualization Techniques	2016
On the Performance of Lossy Compression Schemes for Energy Constrained Sensor Networking	2014
A Multi-agent System with Reinforcement Learning Agents for Biomedical Text Mining	2015
Keynote Talk III: A Formal Methods Perspective on Product Line Engineering	2015
Assessing Model-based Testing: An Empirical Study Conducted in Industry	2014
Monitoring and Information on Skills Development at University: A Multiple-case Study	2016
Analyzing Technology-enhanced Knowledge Practices in an Engineering Course	2009
Improving the Design and Impact of Interactive, Dynamic Visualizations for Science Learning	2008
The Impact of Semi-supervised Clustering on Text Classification	2013
Potential and Pitfalls of Domain-Specific Information Extraction at Web Scale	2016
Supporting Efficient Top-k Queries in Type-ahead Search	2012
Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums	2008

Title	Year
Field Damage Assessments As a Design Tool for Information and Communications Technology Systems That Are Resilient to Natural Disasters	2011
A Survey on Ontologies for Human Behavior Recognition	2014
Towards the Validation of a Method for Quantitative Mobile Usability Testing Based on Desktop Eyetracking	2012
Bi2SoN: A Digital Library for Supporting Biomedical Research	2012
Rapid Business Process Discovery (R-BPD)	2007
Thyroid Disease Diagnosis Using Artificial Immune Recognition System (AIRS)	2009
Development of Indian Weighted Diabetic Risk Score (IWDRS) Using Machine Learning Techniques for Type-2 Diabetes	2016
Adaptive Strategies for Dynamic Pricing Agents	2011
On the Reliability of Profile Matching Across Large Online Social Networks	2015
Streaming Quotient Filter: A Near Optimal Approximate Duplicate Detection Approach for Data Streams	2013
Interaction Design and Cognitive Gameplay: Role of Activation Time	2014
On the Competitive Ratio of Evaluating Priced Functions	2011
Self-adaptation in a Network of Social Drivers: Using Random Boolean Networks	2011
Physical and Digital Media Usage Patterns on Interactive Tabletop Surfaces	2010
Designing a Citation-sensitive Research Tool: An Initial Study of Browsing-specific Information Needs	2009
A Study of Query Reformulation for Patent Prior Art Search with Partial Patent Applications	2015
On the Prediction of Popularity of Trends and Hits for User Generated Videos	2013
Cookies: A Deployment Study and the Testing Implications	2009
SemEval-2007 Task 16: Evaluation of Wide Coverage Knowledge Resources	2007
Plea Against Cultural Stereotypes	2010
Zero-day Malware Detection Based on Supervised Learning Algorithms of API Call Signatures	2011
Understanding the Influence and Service Type of Trusted Third Party on Consumers' Online Trust: Evidence from Australian B2C Marketplace	2016
A Fast Runtime Fault Recovery Approach for NoC-Based MPSoCS for Performance Constrained Applications	2014
Measuring Information Volatility in a Health Care Information Supply Chain	2009
Improving Text Classification Accuracy by Training Label Cleaning	2013
An UML-based Approach to Software Development Cost Estimation	2014
Prediction of Meteorological Parameters: A Semantic Kriging Approach	2016
Design of the Survey Instrument to Evaluate Innovation in Healthcare Centers	2015
Information-driven Interaction-oriented Programming: BSPL, the Blindingly Simple Protocol Language	2011
Evaluation in Education and Guidance	2015
End-to-end Secure Delivery of Scalable Video Streams	2009
Verbal and Nonverbal Communication for Evaluating Interactive Spaces.	2013
Publish/Subscribe Architecture for Mobile Ad Hoc Networks	2008
Information Retrieval Evaluation with Humans in the Loop	2014
Apolo: Making Sense of Large Network Data by Combining Rich User Interaction and Machine Learning	2011
The Effectiveness of Marketing Strategies in Social Media: Evidence from Promotional Events	2015
An Automated Approach to Monitoring and Diagnosing Requirements	2007
Facilitating Parent-teenager Communication Through Interactive Photo Cubes	2013
Classifying Latent User Attributes in Twitter	2010
Modeling User Story Completion of an Agile Software Process	2013
Planning for the Unknown: Lessons Learned from Ten Months of Non-participant Exploratory Observations in the Industry	2015
OSDC: Adapting ODC for Developing More Secure Software	2013
Context Retrieval for Web Tables	2015
Making Cultures: Building Things & Building Communities	2014
Evidence-based Innovation Methodology As a Way to Produce Open Educational Resources by In-service Teachers	2016
Nonnegative Least-Squares Methods for the Classification of High-Dimensional Biological Data	2013
Two's Company, Three's a Crowd: A Case Study of Crowdsourcing Software Development	2014
Software or Wetware?: Discovering when and Why People Use Digital Prosthetic Memory	2007
Identification of Genomic Features in the Classification of Loss- and Gain-of-Function Mutation: [Extended Abstract]	2014
Mining Noisy Tagging from Multi-label Space	2012
4th Annual Workshop on Interoperability and Smart Interactions in Healthcare (ISIH)	2012
Russian Web Spam Evolution: Yandex Experience	2013
Influence Maximization: Near-optimal Time Complexity Meets Practical Efficiency	2014
Understanding Mobile Searcher Attention with Rich Ad Formats	2016
Supporting Ranking Pattern-based Aggregate Queries in Sequence Data Cubes	2009
Identifying Health Domain URLs Using SVM	2015
Variation Aware Optimal Threshold Voltage Computation for On-chip Noise Sensors	2014
An Empirical Analysis of Paid Placement in Online Advertising	2007
Keeping Up with the Tweet-dashians: The Impact of 'Official' Accounts on Online Rumoring	2016
Effective Feedback: Guidelines for Improving Performance	2008
Revisiting Aggregation Techniques for Big Data	2013
Impact of Multimedia Case Studies	2011
OPT: A New Framework for Overlapped and Parallel Triangulation in Large-scale Graphs	2014
Time to Reflect: Supporting Health Services over Time by Focusing on Collaborative Reflection	2016
Letter from the Editors	2008
Managing Complexity in Pre-operative Information Management Systems	2011
Visualizing Dimensionally-reduced Data: Interviews with Analysts and a Characterization of Task Sequences	2014
An Enhanced Test Case Selection Approach for Model-based Testing: An Industrial Case Study	2010
Evaluation of Vision-based Real-time Measures for Emotions Discrimination Under Uncontrolled Conditions	2013
SkylineSearch: Semantic Ranking and Result Visualization for Pubmed	2011
Directing Exploratory Search with Interactive Intent Modeling	2013
On the Protocol Composition Logic PCL	2008
VizioMetrix: A Platform for Analyzing the Visual Information in Big Scholarly Data	2016
Beyond The Thesis: Completing A Successful PhD	2015
Spectral Geometry for Simultaneously Clustering and Ranking Query Search Results	2008

Title	Year
The Effect of Experience on System Usability Scale Ratings	2012
A 24H Forecast of Solar Irradiance Using Echo State Neural Networks	2015
Revisiting the Foundations of IR: Timeless, Yet Timely	2015
High Performance Computing with Spatial Datasets	2010
Capacitated Vehicle Routing System Applying Monte Carlo Methods	2015
Service-oriented Architecture Modeling: Bridging the Gap Between Structure and Behavior	2011
From the Editor	2009
Mining Featured Patterns of MiRNA Interaction Based on Sequence and Structure Similarity	2013
An Entropy-based Framework for Efficient Post-disaster Assessment Based on Crowdsourced Data	2016
Stock Risk Mining by News	2010
Investigating Collaborative Learning Success with Physiological Coupling Indices Based on Electrodermal Activity	2016
Social Glass: A Platform for Urban Analytics and Decision-making Through Heterogeneous Social Data	2015
A Study of Selective Collection Enrichment for Enterprise Search	2009
Parallel Optimal Pairwise Biological Sequence Comparison: Algorithms, Platforms, and Classification	2016
An Empirical Investigation into the Role of API-level Refactorings During Software Evolution	2011
TAF: Temporal Analysis Framework for Handling Data Velocity in Healthcare Analytics	2016
Using an Ant Colony Optimization Algorithm for Monotonic Regression Rule Discovery	2016
DrugNerAR: Linguistic Rule-based Anaphora Resolver for Drug-drug Interaction Extraction in Pharmacological Documents	2009
Health Problem Solving by Older Persons Using a Complex Government Web Site: Analysis and Implications for Web Design	2011
Work Practices of System Administrators: Implications for Tool Design	2008
How to Protect Adjacent Bridges Due to Tunnel Excavation: A Case Study	2015
A PSO/ACO Approach to Knowledge Discovery in a Pharmacovigilance Context	2009
On the Effectiveness of API-level Access Control Using Bytecode Rewriting in Android	2013
Globally Distributed System Developers: Their Trust Expectations and Processes	2013
CLUES: A Unified Framework Supporting Interactive Exploration of Density-based Clusters in Streams	2011
Probabilistic Ranked Queries in Uncertain Databases	2008
Agile Security Issues: An Empirical Study	2010
The Whys, How Tos, and Pitfalls of User Studies	2009
Engagement-aware Computing: Modelling User Engagement from Mobile Contexts	2016
A Constraint-based Approach to Guarded Algebraic Data Types	2007
Understanding the Relationship Between Frustration and the Severity of Usability Problems: What Can Psychophysiological Data (Not) Tell Us?	2016
Enhancing Knowledge Base with Knowledge Transfer	2012
The Research Trends and Comparing About the Big Data Between Korea and China Using Text Mining	2015
Resource Optimization for Processing of Stream Data in Data Warehouse Environment	2012
Outlier Ensembles: Position Paper	2013
Evaluating Enjoyment Within Alternate Reality Games	2011
2Nd Software Certification Consortium Workshop: Theoretical Basis for System and Software Engineering Practices and Certification	2011
Challenges and Opportunities for Trust Management in Crowdsourcing	2012
ICT Practices by Voluntary Groups: A Multi-setting Study	2013
Arctic Ice, George Clooney, Lipstick on a Pig, and Insomniac Fruit Flies: Combining Kd and M&s for Predictive Analysis	2011
Effective Deep Learning-based Multi-modal Retrieval	2016
Identification of Information Requirements Using Simulation for Supporiting Construction Productivity Assessment	2010
Efficiently Answering Top-k Typicality Queries on Large Databases	2007
Accurate Numerical Integration of Perturbed Oscillatory Systems in Two Frequencies	2009
Agent-based Simulation for Assessing Network Security Risk Due to Unauthorized Hardware	2015
A Matter of Time - Intrinsic or Extrinsic - for Diffusion in Evolving Complex Networks	2013
The C-ND Tree: A Multidimensional Index for Hybrid Continuous and Non-ordered Discrete Data Spaces	2009
Personas is Applicable: A Study on the Use of Personas in Denmark	2014
Critical Skills for Game Developers: An Analysis of Skills Sought by Industry	2008
Late Fusion of Heterogeneous Methods for Multimedia Image Retrieval	2008
Secure Containers in Android: The Samsung KNOX Case Study	2016
Designing a Multi-slate Reading Environment to Support Active Reading Activities	2012
A Pattern System for Tracing Architectural Concerns	2011
Managing Security and Privacy in Ubiquitous eHealth Information Interchange	2011
On a New SDP-SOCP Method for Acoustic Source Localization Problem	2016
An Open Approach Towards the Benchmarking of Table Structure Recognition Systems	2010
Ontology Extraction from Stories: An Exploratory Study in Storytelling	2015
Assessor Error in Stratified Evaluation	2010
Design of a Fuzzy Expert System for Determination of Coronary Heart Disease Risk	2007
Security Requirements Embedded in MS Programs in Information Sciences and Technologies	2016
SharpLudus Revisited: From Ad Hoc and Monolithic Digital Game DSLs to Effectively Customized DSM Approaches	2011
Textual Analysis for Studying Chinese Historical Documents and Literary Novels	2015
A Synthesized Definition of Computer Ethics	2014
Unified Tests for Fine Scale Mapping and Identifying Sparse High-dimensional Sequence Associations	2015
A Testbed for Adaptive Security for IoT in eHealth	2013
A Non-negative Matrix Tri-factorization Approach to Sentiment Classification with Lexical Prior Knowledge	2009
Dependency-Based Construction of Semantic Space Models	2007
Governance of IT in the Network Society	2010
Using Virtual Environments to Assess Time-to-contact Judgments from Pedestrian Viewpoints	2007
Absolute and Relative Clustering	2013
Organizational Volatility and Its Effects on Software Defects	2010
Developer Fluency: Achieving True Mastery in Software Projects	2010
TraClass: Trajectory Classification Using Hierarchical Region-based and Trajectory-based Clustering	2008
How Low Can You Go?: Is the System Usability Scale Range Restricted?	2013
Random Testing of Purely Functional Abstract Datatypes: Guidelines for Dealing with Operation Invariance	2013

Title	Year
Search Result Diversification	2010
Strong Mutation-based Test Data Generation Using Hill Climbing	2016
A Semantics Based Symbolic Regression Framework for Mining Explicit and Implicit Equations from Data	2016
User Behavior Driven Ranking Without Editorial Judgments	2010
Leveraging Candidate Popularity on Twitter to Predict Election Outcome	2013
Grouping Game Players Using Parallelized K-means on Supercomputers	2015
Stay or Quit: IT Personnel Turnover in Botswana	2011
Anti-social Media: Communicating Risk Through Open Data, Crime Maps and Locative Media	2014
Analysis in Usability Evaluations: An Exploratory Study	2010
Mining Linguistic Cues for Query Expansion: Applications to Drug Interaction Search	2009
Divide and Recycle: Types and Compilation for a Hybrid Synchronous Language	2011
Reliability Prediction of Webpages in the Medical Domain	2012
Expanding Network Communities from Representative Examples	2009
Solving Nuclear Safeguards Evaluation Problem with Fuzzy Multiple Attribute Decision Making Methods	2009
Reducing the Footprint of Certifiable Health Software During Early Stage Development	2011
An Integrated Approach to Information Systems Security Policy Violation: The Case of Ethiopia	2016
Reverse Engineering the Amen Break: Score-informed Separation and Restoration Applied to Drum Recordings	2016
Validation of a Learning Progression: Relating Empirical Data to Theory	2010
Musiplectics: Computational Assessment of the Complexity of Music Scores	2015
Querylog-based Assessment of Retrievability Bias in a Large Newspaper Corpus	2016
Visual Recognition in Museum Guide Apps: Do Visitors Want It?	2014
Instance Selection for Machine Translation Using Feature Decay Algorithms	2011
Synthesizing High Utility Suggestions for Rare Web Search Queries	2011
A Preference Approach to Reputation in Sponsored Search	2016
Extracting Domain Models from Natural-language Requirements: Approach and Industrial Evaluation	2016
Exploring Misconceptions of Operating Systems in an Online Course	2013
Querying K-truss Community in Large and Dynamic Graphs	2014
Why CSCW Needs Science Policy (and Vice Versa)	2013
Cyber-physical System Product Line Engineering: Comprehensive Domain Analysis and Experience Report	2015
ICGA-PSO-ELM Approach for Accurate Multiclass Cancer Classification Resulting in Reduced Gene Sets in Which Genes Encoding Secreted Proteins Are Highly Represented	2011
A Histogram-based Technique for Automatic Threshold Assessment in a Run Length Smoothing-based Algorithm	2010
Experiments with a Venue-Centric Model for Personalisedand Time-Aware Venue Suggestion	2015
Integrating the Visualisation Reference Model with Ecological Interface Design	2007
Some Performance Bounds of Strategies for Graph Exploration	2015
A Quantitative Study of Accuracy in System Call-based Malware Detection	2012
Answering Bounded Continuous Search Queries in the World Wide Web	2007
Where Are You Heading, Metric Access Methods?: A Provocative Survey	2010
Consumer Valuations on Digital Product Innovation	2016
Improved Indexing & Searching Throughput	2016
A Guide to Reporting Scientific Evaluation in Visualization	2012
Exploring the African Village Metaphor for Computer User Interface Icons	2009
Sympathetic Guitar: Humans Respond Socially to Interactive Technology in an Abstract, Expressive Context	2011
Semantic Convergence of Wikipedia Articles	2007
An Approach to Spanish Primary School Teachers' Attitudes Towards Collaborative Learning with Video Games and the Influence of Teacher Training	2016
Digital Humanities Development Without Developers: Bulgarian Dialectology As Living Tradition	2014
A Data-oriented Survey of Context Models	2007
Quick Sensitivity Analysis for Incremental Data Modification and Its Application to Leave-one-out CV in Linear Classification Problems	2015
Correlational Analysis Between School Performance and Municipal Indicators in Brazil Supported by Linked Open Data	2016
Evolution by Adapting Surrogates	2013
Scalable and Axiomatic Ranking of Network Role Similarity	2014
Object-Oriented Modelling and Simulation of Air Flow in Data Centres Based on a Quasi-3D Approach for Energy Optimisation	2014
Towards Cross-platform Cross-language Analysis with Soot	2016
ParTime: Parallel Temporal Aggregation	2016
TAGs: Scalable Threshold-based Algorithms for Proximity Computation in Graphs	2011
Keyword Search in Databases: The Power of RDBMS	2009
Large-scale Multimodal Semantic Concept Detection for Consumer Video	2007
The Simulation Life-cycle: Supporting the Data Collection and Representation Phase	2014
Citation and Attribution in Open Science: A Case Study	2016
Exploring QoE for Power Efficiency: A Field Study on Mobile Videos with LCD Displays	2015
Harnessing the Quantified Self Movement for Optimal Mental Health and Wellbeing	2016
Energy Advisors at Work: Charity Work Practices to Support People in Fuel Poverty	2014
Archetypal Networks	2015
Systems-of-systems Framework for Providing Real-time Patient Monitoring and Care	2014
Rapid Deployment of Simulation Models for Building Construction Applications	2009
Prediction of Meteorological Parameters: An A-posteriori Probabilistic Semantic Kriging Approach	2016
Qualitative Assessment of Contemporary Media Sharing Practices and Their Relationship to the sMS Platform	2012
Theoretical Underpinnings of Computing Education Research: What is the Evidence?	2014
MTSS – A Marine Traffic Simulation System and Scenario Studies for a Major Hub Port	2016
D-Sieve: A Novel Data Processing Engine for Efficient Handling of Crises-Related Social Messages	2015
Large-scale Validation and Analysis of Interleaved Search Evaluation	2012
Design Parameters of Rating Scales for Web Sites	2007
An Algorithm for Selective Undo of Any Operation in Collaborative Applications	2010
Social Media and Online Survey: Tools for Knowledge Management in Health Research	2014
Cross-workplace Perspectives: Relating Studies from Hospitals to an Oil and Gas Workplace	2008

Title	Year
The Impact of Random Samples in Ensemble Classifiers	2010
Exploring the Use of Wikis for Information Sharing in Interdisciplinary Design	2009
Knowlywood: Mining Activity Knowledge From Hollywood Narratives	2015
Remote Device Support in Thin Client Network	2010
Synthetic Biology and Microdevices: A Powerful Combination	2013
Facilitating Synchronization and Coordination Within Dispersed Emergency Management Teams	2008
Modelling for the Triple-bottom Line: An Investigation of Hybrid Simulation for Sustainable Development Analysis	2016
Modelling for the Triple-bottom Line: An Investigation of Hybrid Simulation for Sustainable Development Analysis	2016
Level-set-based Partitioning and Packing Optimization of a Printable Model	2015
Large Scale Image Copy Detection Evaluation	2008
Preliminary Study of Stress/Neutral Detection on Recordings of Children in the Natural Home Environment	2009
On the Lifetime of Wireless Sensor Networks	2009
An Online Approach to Dynamic Channel Access and Transmission Scheduling	2015
Mutation Operators for JADE Mobile Agent Systems	2012
Factorized Sparse Learning Models with Interpretable High Order Feature Interactions	2014
Behavioral Game Theoretic Models: A Bayesian Framework for Parameter Analysis	2012
Best Practices for Web Applications to Improve Performance of QoS	2016
Swiping Paper: The Second Hand, Mundane Artifacts, Gesture and Collaboration	2010
Measuring Coherence Between Electronic and Manual Annotations in Biological Databases	2009
Quantifying Deception: A Case Study in the Evolution of Antimicrobial Resistance	2016
PROTON: An Automatic Place-and-route Tool for Optical Networks-on-chip	2013
Linking Stanford Typed Dependencies to Support Text Analytics	2015
A Probabilistic Topic-based Ranking Framework for Location-sensitive Domain Information Retrieval	2009
Parallel Data Cubes on Multi-core Processors with Multiple Disks	2011
A Fast and Practical Approach to Genotype Phasing and Imputation on a Pedigree with Erroneous and Incomplete Information	2012
Learning Biological Networks via Bootstrapping with Optimized Go-based Gene Similarity	2010
A Localized and Distributed Channel Assignment Framework for Cognitive Radio Networks	2007
Discretionary Social Network Data Revelation with a User-centric Utility Guarantee	2012
On Supporting Obstacle-free and Sharing Tree for Information Collection in WSNs	2008
Performance-aware Cross-layer Design in Wireless Multihop Networks via a Weighted Backpressure Approach	2016
Organizational Commitment of IT Workers: Leader Support and Differences Across Gender and Race	2013
Recognizing Biomedical Named Entities Using Skip-chain Conditional Random Fields	2010
Outlier Ensembles	2013
Analysis Problems for Graphical Dynamical Systems: A Unified Approach Through Graph Predicates	2015
A Hybrid Model for Task Completion Effort Estimation	2016
Re-engineering Relational Databases: The Way Forward	2011
Molecular Dynamics Simulations on High-Performance Reconfigurable Computing Systems	2010
Effective Keyword-based Selection of Relational Databases	2007
Population Restarting: A Study Case of Feature Extraction from Remotely Sensed Imagery Using Textural Information	2010
Citations and Annotations in Classics: Old Problems and New Perspectives	2013
Developing a Questionnaire for Measuring Mobile Business Service Experience	2008
Designing for Citizen Data Analysis: A Cross-Sectional Case Study of a Multi-Domain Citizen Science Platform	2015
Relative Status of Journal and Conference Publications in Computer Science	2010
Regression on Manifolds Using Kernel Dimension Reduction	2007
Spy vs. Spy: Counter-intelligence Methods for Backtracking Malicious Intrusions	2011
A Supervised Learning Approach to Detect Subsumption Relations Between Tags in Folksonomies	2015
DIRECT: A Scalable Approach for Route Guidance in Selfish Orienteering Problems	2015
Theory and Algorithms for Hop-count-based Localization with Random Geometric Graph Models of Dense Sensor Networks	2012
Key Factors for Adopting Inner Source	2014
A Query Substitution-Search Result Refinement Approach for Long Query Web Searches	2009
So Many Topics, So Little Time	2009
Efficient Pruning Algorithm for top-K Ranking on Dataset with Value Uncertainty	2013
PressureFish: A Method to Improve Control of Discrete Pressure-based Input	2008
Theorem-based, Data-driven, Cyber Event Detection	2013
Increasing Robustness of Multimodal Interaction via Individual Interaction Histories	2016
A UML Profile for Dependability Analysis of Real-time Embedded Systems	2007
Case Study: Designing an Advanced Visualization System for Geological Core Drilling Expeditions	2010
Finding Answers in Web Search	2015
Craving, Creating, and Constructing Comfort: Insights and Opportunities for Technology in Hospice	2014
Establishing Safety Case Strategies for Mission Planning or Situational Awareness Systems	2011
Sketching Sketching: Outlines of a Collaborative Design Method	2009
On Learning-based Methods for Design-space Exploration with High-level Synthesis	2013
The Shortest Path: Comparison of Different Approaches and Implementations for the Automatic Routing of Vehicles	2012
A Novel Layout Decomposition Algorithm for Triple Patterning Lithography	2012
Second Life for Illiterates: A 3D Virtual World Platform for Adult Basic Education	2010
Verification of Data-aware Commitment-based Multiagent System	2014
Robots Against Infectious Diseases	2016
Social Negative Bootstrapping for Visual Categorization	2011
Routines That Ease the Pain: The Information World of a Dialysis Clinic	2010
Do Software Libraries Evolve Differently Than Applications?: An Empirical Investigation	2007
Methods of Working with Teenagers in Interaction Design	2013
Baseline Evaluation: An Empirical Study of the Performance of Machine Learning Algorithms in Short Snippet Sentiment Analysis	2014
Software Project Managers' Perceptions of Productivity Factors: Findings from a Qualitative Study	2016
Games with a Purpose or Mechanised Labour?: A Comparative Study	2013
A Knowledge-based Database System for Visual Rating of fMRI Activation Patterns for Brain Language Networks	2009

Title	Year
The Effect of Synonym Substitution on Search Results	2016
Experimental Evaluation and Hypertexts	2011
Inferring Topic-dependent Influence Roles of Twitter Users	2014
Spotting Working Code Examples	2014
3rd Annual Workshop on Interoperability and Smart Interactions in Healthcare (ISIH)	2011
Online Reorganization of Databases	2009
Selectivity-based XML Query Processing in Structured Peer-to-peer Networks	2010
Towards a Better Understanding of Group Forking Dynamics in Virtual Contexts	2009
Interactive Record/Replay for Web Application Debugging	2013
Heterogeneity of IT Employees: An Analysis of a Model of Perceived Organizational Support by Job Type	2011
Studying Volatility Predictors in Open Source Software	2012
Evolving e-Government Benchmarking to Better Cover Technology Development and Emerging Societal Needs	2012
PruSM: A Prudent Schema Matching Approach for Web Forms	2010
A Comparative Study of Minimally Supervised Morphological Segmentation	2016
The Co-evolution of Communities of Practice and Knowledge Management in Organizations	2012
PianoText: Redesigning the Piano Keyboard for Text Entry	2014
'Is' to 'Was': Coordination and Commemoration in Posthumous Activity on Wikipedia Biographies	2015
Inflow and Retention in OSS Communities with Commercial Involvement: A Case Study of Three Hybrid Projects	2016
Firework Visualization: A Model for Local Citation Analysis	2012
The Ironies of Automation: Still Going Strong at 30?	2012
Measuring Productivity in Agile Software Development Process: A Scoping Study	2015
Simulation by Example for Complex Systems	2014
Comparing the Performance of Fault Prediction Models Which Report Multiple Performance Measures: Recomputing the Confusion Matrix	2012
Recovering Temporally Rewiring Networks: A Model-based Approach	2007
Learning at Scale: Using an Evidence Hub To Make Sense of What We Know	2016
Optimally Discriminative Choice Sets in Discrete Choice Models: Application to Data-Driven Test Design	2016
Identifying Influential Bloggers: Time Does Matter	2009
Spatial Trajectories Segmentation: Trends and Challenges	2016
Uncertain Time-series Similarity: Return to the Basics	2012
Requirements Attributes to Predict Requirements Related Defects	2010
Sensing Stress Network for Social Coping	2014
A Pattern Language for Use Cases Specification	2015
The Effects of the Reliability of an Automatic Target Recognition System on Image Analyst Performance	2007
A Standard-driven Implementaion of WS-BPEL 2.0	2010
Dynamic Data-flow Testing	2014
SIG: Gender-Inclusive Software: What We Know About Building It	2015
On the Validity of Flow-level TCP Network Models for Grid and Cloud Simulations	2013
The Impact of Co-evolution of Dynamic Networks Upon Adolescent Deviant Behaviors	2015
AJAR: Aggregations and Joins over Annotated Relations	2016
Cost-Effective Traceability Links for Architecture-Level Software Understanding: A Controlled Experiment	2015
The Exact Distance to Destination in Undirected World	2012
Systems Aspects of Probabilistic Data Management	2008
Advanced Probabilistic Couplings for Differential Privacy	2016
On the Relative Importance of Individual Components of Chord Recognition Systems	2014
DEFCAM: A Design and Evaluation Framework for Defect-tolerant Cache Memories	2011
A System for Interactive Visualization of Off-screen Objects on Mobile Devices	2012
Diversity-oriented Multimodal and Interactive Information Retrieval	2016
Proposition and Validation of a New Index to Determine the Measurement Change Resolution of Inertial Motion Tracking Systems	2016
Social Groups, Social Media, and Civic Participation of High School Youth: Concepts and Methods for Design Implications	2009
Semantically Faceted Navigation with Topic Pies	2014
Understanding Energy Consumption at Work: Learning from Arrow Hill	2016
The Strength of Awkward Ties: Online Interactions Between High School Students and Adults	2016
Blueprints of a Lightweight Automated Experimentation System: A Building Block Towards Experimental Cyber Security	2011
Traffic Analysis of Mobile Broadband Networks	2007
Fragmentation and Choreography: Caring for a Patient and a Chart During Childbirth	2012
Social Question Answering: Textual, User, and Network Features for Best Answer Prediction	2016
Inferring International and Internal Migration Patterns from Twitter Data	2014
Indexing Correlated Probabilistic Databases	2009
Dependency Oriented Complexity Metrics to Detect Rippling Related Design Defects	2009
A Simulation-based Approach to Statistical Inventory Management	2012
A Survey of Provably Secure Searchable Encryption	2014
Multimodal Deception Detection: A T-pattern Approach	2015
Apples-to-apples in Cross-validation Studies: Pitfalls in Classifier Performance Measurement	2010
Process Culture: Maturity, Drivers, and Barriers	2016
Component Airbag: A Novel Approach to Develop Dependable Component-based Applications	2007
Decentralized Human Milk Banking with ODK Sensors	2013
Time Series Forecasting Using a Perturbative Intelligent System	2010
Continuous Cotemporal Probabilistic Modeling of Systems Biology Networks from Sparse Data	2011
Introducing Digital Logic and Electronics Concepts in a Game-like Environment	2009
Software Development Risk Management: Using Machine Learning for Generating Risk Prompts	2015
Adaptive User Feedback for IR-based Traceability Recovery	2015
Progressive Processing of Subspace Dominating Queries	2011
Teaching Game Design and Game Programming Through Interdisciplinary Courses	2008
On the Failure of Game Theoretic Approach for Distributed Deadlock Resolution	2012
Tracking Web Spam with HTML Style Similarities	2008

Title	Year
At a Different Tempo: What Goes Wrong in Online Cross-cultural Group Chat?	2012
BRAVE: Bit-rate Adaptation in Vehicular Environments	2012
Software Process Models vs. Descriptions: What Do Practitioners Use and Need?	2016
Study of the Electromagnetic Scattering Models for Water Parameters Estimation from TerraSAR-X Images	2016
Beyond Classical Measures: How to Evaluate the Effectiveness of Interactive Information Retrieval System?	2007
Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit	2015
Visual Menu Techniques	2016
Dynamic Networks Reveal Key Players in Aging	2013
Reusing Debugging Knowledge via Trace-based Bug Search	2012
Reusing Debugging Knowledge via Trace-based Bug Search	2012
Look Ma, No Email!: Blogs and IRC As Primary and Preferred Communication Tools in a Distributed Firm	2011
How Robot Verbal Feedback Can Improve Team Performance in Human-Robot Task Collaborations	2015
Searching Online Book Documents and Analyzing Book Citations	2013
Differences and Similarities Between Finger and Pen Stroke Gestures on Stationary and Mobile Devices	2015
On the Design and Evaluation of Producer Mobility Management Schemes in Named Data Networks	2015
Improved Approximations for the Hotlink Assignment Problem	2011
A Methodology to Derive Sustainability Indicators for Software Development Projects	2013
Late Fusion of Compact Composite Descriptors for Retrieval from Heterogeneous Image Databases	2010
Understanding HCI Education Across Asia-Pacific	2015
Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don'T Know About the As Graph	2008
A Comprehensive Approach to Trust Management	2013
The Relationship Between Problem Frequency and Problem Severity in Usability Evaluations	2014
Towards a Complexity Theory for Local Distributed Computing	2013
Challenges in Software Ecosystems Research	2015
SimDB: A Similarity-aware Database System	2010
The Motivation and Proposition of a Privacy-enhancing Architecture for Operational Databases	2007
Modelling and Analysis of the Split Phase Multi-channel MAC Approach	2012
Survival Needs and Social Inclusion: Technology Use Among the Homeless	2010
Towards Automatic Extraction of Epistemic Items from Scientific Publications	2010
A Case Study into the Accessibility of Text-parser Based Interaction	2015
Modeling the Communication Costs of Content-based Routing: The Case of Subscription Forwarding	2007
Design of a Health Monitoring Toy for Children	2016
Examining Life at the Code Face	2008
A Surrogate Competition Approach to Enhancing Game-based Learning	2013
Florch: Challenges on Developing a New Social Network Accessible for Senescent Users	2015
Digital Mobilization in Disaster Response: The Work & Self-organization of On-line Pet Advocates in Response to Hurricane Sandy	2014
Scalable and Efficient Web Search Result Diversification	2016
A Study of Partitioning and Parallel UDF Execution with the SAP HANA Database	2014
Panel Session: Discourse Annotation	2007
POSTER: Security Enhanced Administrative Role Based Access Control Models	2016
What Do Teens Ask Their Online Social Networks?: Social Search Practices Among High School Students	2014
The Complexity of Perception of Image Distortion: An Initial Study	2010
Examining the Impact of Regional Familiarity and Design Characteristics on Use of a Map-based News Visualization	2013
Knowing Where and when to Look in a Time-critical Multimodal Dual Task	2010
RSA-MBT: A Test Tool for Generating Test Artifacts Based on Models	2013
Consumer Trust in e-Commerce Web Sites: A Meta-study	2011
Approximation Algorithms for Speeding Up Dynamic Programming and Denoising aCGH Data	2011
End-user Debugging Strategies: A Sensemaking Perspective	2012
A Call for Energy Efficiency in Data Centers	2014
Characterization & Analysis of a Server Consolidation Benchmark	2008
Evaluating Enjoyment Within Alternate Reality Games	2011
Modeling Tagged Photos for Automatic Image Annotation	2011
What's It Worth to You?: The Costs and Affordances of CMC Tools to Asian and American Users	2010
Functional Size Measurement Revisited	2008
A Study of Hierarchical and Flat Classification of Proteins	2010
Towards an Axiomatization of Statistical Privacy and Utility	2010
Safe: Simulation Automation Framework for Experiments	2012
AuDroid: Preventing Attacks on Audio Channels in Mobile Devices	2015
Interactive Dancing Game with Real-time Recognition of Continuous Dance Moves from 3D Human Motion Capture	2011
Multiscale Modeling and Analysis of Planar Cell Polarity in the Drosophila Wing	2013
Adaptive Contract Design for Crowdsourcing Markets: Bandit Algorithms for Repeated Principal-agent Problems	2014
The CareRabbit	2010
Robust Gender Classification on Unconstrained Face Images	2015
Modeling Play: Re-casting Expertise in MMOGs	2011
Individual and Aggregate Electrical Load Forecasting: One for All and All for One	2015
Performance Modelling and Analysis of Software-Defined Networking Under Bursty Multimedia Traffic	2016
Evaluation of Robust Estimation Methods in Estimating Weibull Parameters for Wind Energy Application	2015
Distance-based Outlier Detection: Consolidation and Renewed Bearing	2010
User-centric Inference Based on History of Context Data in Pervasive Environments	2008
Modeling Play: Re-casting Expertise in MMOGs	2011
Facts or Friends?: Distinguishing Informational and Conversational Questions in Social Q&A Sites	2009
How to Design Educational Technologies?: The Development of an Instructional-model	2008
A Survey and Taxonomy Aimed at the Detection and Measurement of Covert Channels	2016
Evaluating the Use of Learning Objects in CS1	2011
Comparing Approaches for Query Autocompletion	2015
A Comparative Case Study on Clients Participation in a 'Traditional' and in an Agile Software Company	2011

Title	Year
Complexity Measures for Secure Service-Oriented Software Architectures	2007
Lost in Navigation: Evaluating a Mobile Map App for a Fair	2012
EyeGaze: Enabling Eye Contact over Video	2014
From the Eye to the Heart: Eye Contact Triggers Emotion Simulation	2012
Computational Methods for Residential Energy Cost Optimization in Smart Grids: A Survey	2016
Broadcasting in Multihop Mobile Tactical Networks: To Network Code or Not	2010
A Multi-spectral Stereo Method to Retrieve Cloud Top Height Applied to Geostationary Satellite Images	2016
Evaluating Software Specifications by Comparison	2011
Visual Analysis of Time-motion in Basketball Games	2010
From Interaction Science to Cognitive Interaction Technology	2014
Online Planning for Large MDPs with MAXQ Decomposition	2012
Designing for the Other 'Hereafter': When Older Adults Remember About Forgetting	2016
Multi-User Usability Guidelines for Interactive Wall Display Applications	2015
Integrating Informatics in Humanities Activities and Research	2010
Defining Standards for Web Page Performance in Business Applications	2015
Using Mobile Computing to Support Malnutrition Management in South Africa	2014
Reuse and Optimization of Testbenches and Properties in a TLM-to-RTL Design Flow	2008
Instruction Offloading with HMC 2.0 Standard: A Case Study for Graph Traversals	2015
Towards Arabic Semantic Opinion Mining: Identifying Opinion, Polarity and Intensity	2016
A Survey on Energy Efficiency in P2P Systems: File Distribution, Content Streaming, and Epidemics	2015
Accelerating MATLAB Image Processing Toolbox Functions on GPUs	2010
Using Messy Genetic Algorithms for Solving the Winner Determination Problem	2010
Investigating Sighted Users' Browsing Behaviour to Assist Web Accessibility	2008
Distributed-fountain Network Code (DFNC) for Content Delivery in Vehicular Networks	2013
Evaluation of Disambiguation Mechanisms of Object-Based Selection in Virtual Environment: Which Performances and Features to Support Pick Out?	2014
Improving Biomedical Document Retrieval Using Domain Knowledge	2008
Diagnosing Learners' Problem-Solving Strategies Using Learning Environments with Algorithmic Problems in Secondary Education	2009
BMExpert: Mining MEDLINE for Finding Experts in Biomedical Domains Based on Language Model	2015
A Measurement Study of Interference Modeling and Scheduling in Low-power Wireless Networks	2008
User-Oriented Evaluation Methods for Interactive Web Search Interfaces	2007
Theories and Practice of Design for Information Systems: Eight Design Perspectives in Ten Short Weeks	2008
Chinese Semantic Role Labeling with Shallow Parsing	2009
Don'T Turn Social Media into Another 'Literary Digest' Poll	2011
Logging Every Footstep: Quantile Summaries for the Entire History	2010
Software Engineering Principles: A Survey and an Analysis	2010
Content Replication and Delivery in Vehicular Networks	2014
An Agent-Based Modeling Approach to Investigate Emergent Patterns in Ecological Systems	2010
Temporal Click Model for Sponsored Search	2010
A Comprehensive Survey of Software Development Cost Estimation Studies	2015
Locomotion by Natural Gestures for Immersive Virtual Environments	2016
Automatic Metric Thresholds Derivation for Code Smell Detection	2015
Are All Bits Equal?: Experimental Study of IEEE 802.11 Communication Bit Errors	2012
Estimating Software-intensive Projects in the Absence of Historical Data	2013
Profiling of Symmetric-encryption Algorithms for a Novel Biomedical-implant Architecture	2008
A Hybrid Decision Tree Based Methodology for Event Studies and Its Application to e-Commerce Initiative Announcements	2012
A New Methodology for Diagnosis of Appendicitis Using Sonographic Images in Image Mining	2012
Modeling User Behavior Using a Search-engine	2007
Multiple Inheritance in AgentSpeak(L)-Style Programming Languages	2014
Application of Matrix Completion on Water Treatment Data	2015
Using Web Technologies for Integrative Drug Discovery	2010
Comparing the Use of Discrete-event Simulation and System Dynamics Models	2007
Automatic Extraction of Petrographic Features from Pottery of Archaeological Interest	2015
Adding Individual Patient Case Data to the Melanoma Targeted Therapy Advisor	2013
DICLENS: Divisive Clustering Ensemble with Automatic Cluster Number	2012
Period Optimization for Hard Real-time Distributed Automotive Systems	2007
Exploiting Concurrency for the Automated Synthesis of MPSoC Interconnects	2015
Empirical Analysis of the Performance of Variance Estimators in Sequential Single-run Ranking & Selection: The Case of Time Dilation Algorithm	2016
Bangla Online Handwriting Recognition Using Recurrent Neural Network Architecture	2016
An Experimental Analysis of the Performance of SideChain Packing Algorithms	2015
Semantics-aware Open Information Extraction in the Biomedical Domain	2012
Traffic Prediction and QoS Transmission of Real-time Live VBR Videos in WLANs	2011
Positioning of Media Economic Framework for Novel Media in the Field of Project Business	2011
Model-free Variable Selection in Reproducing Kernel Hilbert Space	2016
Utility-based Forwarding: A Comparison in Different Mobility Scenarios	2012
DIGNiFI: Discovering Causative Genes for Orphan Diseases Using Protein-protein Interaction Networks	2016
The Self-feeding Process: A Unifying Model for Communication Dynamics in the Web	2013
Digital Badges for Deliberate Practice: Designing Effective Badging Systems for Interactive Communication Scenarios	2015
Shine: Search Heterogeneous Interrelated Entities	2007
GoPoMoSA: A Goal-oriented Process Modeling and Simulation Advisor	2011
On Approximation Metrics for Linear Temporal Model-checking of Stochastic Systems	2014
Summarization and Matching of Density-based Clusters in Streaming Environments	2011
Mohawk+T: Efficient Analysis of Administrative Temporal Role-Based Access Control (ATRBAC) Policies	2015
Sketches for Size of Join Estimation	2008
Optimal Control of Wireless Networks with Finite Buffers	2012

Title	Year
Impact of Spontaneous Speech Features on Business Concept Detection: A Study of Call-centre Data.	2010
Study of the Reliability of Statistical Timing Analysis for Real-time Systems	2015
StarLight: Exploring Embodied Interactions with Media Architecture and Public Audiences	2015
Incorporating 3D Technologies to the Brazilian DTV Standard: A Study of Integration Strategies Based on Middleware Ginga	2010
Convergence Results for Ant Routing Algorithms Viastochastic Approximation	2010
Graph-based Modelling of Query Sets for Differential Privacy	2016
Sensor Network Localization, Euclidean Distance Matrix Completions, and Graph Realization	2008
Outlier Detection by Logic Programming	2007
Architecting Cloud Tools Using Software Product Line Techniques: An Exploratory Study	2015
Steal Tree: Low-overhead Tracing of Work Stealing Schedulers	2013
Steal Tree: Low-overhead Tracing of Work Stealing Schedulers	2013
Fully-functional Succinct Trees	2010
Distributed Search over Encrypted Big Data	2015
Computational Visual Attention Systems and Their Cognitive Foundations: A Survey	2010
Design Patterns for Developing Dynamically Adaptive Systems	2010
Dependent Interoperability	2012
Study of Robustness of Community Emerged from Exchanges in Networks Communication	2011
Developing a PBL-based Rescue Robotics Course	2009
Push- and Pull-based Epidemic Spreading in Networks: Thresholds and Deeper Insights	2012
Visual and Linguistic Information in Gesture Classification	2007
Markov Chain Monte Carlo Simulation of Biomonitoring in Humans: Application to Biomarkers of Chronic Exposure to Alkyl Benzenes in the Environment	2008
Critical Design and Critical Theory: The Challenge of Designing for Provocation	2012
Exception Handling Negligence Due to Intra-individual Goal Conflicts	2009
Multi-resolution Fourier-based Texture Image Retrieval	2009
Design and Development of a Genetic Algorithm for the Distance Constrained Vehicle Routing Problem with Environmental Issues: Genetic Algorithm for the Green Distance Constrained Vehicle Routing Problem	2015
Dynamic Networks Reveal Key Players in Aging	2014
Remote Navigation of a Mobile Robot in an RFID-augmented Environment	2010
A Buyer-Friendly and Mediated Watermarking Protocol for Web Context	2016
Enabling Information Retrieval on Historical Document Collections: The Role of Matching Procedures and Special Lexica	2009
Maximizing Bichromatic Reverse Nearest Neighbor for Lp-norm in Two- and Three-dimensional Spaces	2011
Main Memory Adaptive Indexing for Multi-core Systems	2014
Using a Soft Systems Methodology Framework to Guide the Conceptual Modeling Process in Discrete Event Simulation	2012
Efficient Two-dimensional Data Allocation in IEEE 802.16 OFDMA	2014
A Taxonomy and Survey of Self-protecting Software Systems	2012
Planning Interventions in Biological Networks	2010
An Experimental Evaluation of Real-time DVFS Scheduling Algorithms	2012
Graph Construction and B-matching for Semi-supervised Learning	2009
Primal or Dual: Which Promises Faster Spatiotemporal Search?	2008
Modeling DVFS and Power-Gating Actuators for Cycle-Accurate NoC-Based Simulators	2015
Fully Functional Static and Dynamic Succinct Trees	2014
Towards a Framework for Migrating Web Applications to Web Services	2011
Cloud Elasticity: Going Beyond Demand As User Load	2016
Nugget Discovery in Visual Exploration Environments by Query Consolidation	2007
A Theoretical Analysis of the Prodrug Delivery System for Treating Antibiotic-Resistant Bacteria	2011
The Ultimate Steganalysis Benchmark?	2007
Situating Cyberinfrastructure in the Public Realm: The TeraGrid and XSEDE Projects	2016
Modeling Best-effort and FEC Streaming of Scalable Video in Lossy Network Channels	2007
Quantitative Analysis of Evolvability Using Vertex Centralities in Phenotype Network	2016
Statistical Analysis of Sketch Estimators	2007
Version History, Similar Report, and Structure: Putting Them Together for Improved Bug Localization	2014
A Survey of Intrusion Detection Techniques for Cyber-physical Systems	2014
Making Computing Interesting to School Students: Teachers' Perspectives	2013
Inspect, Embody, Invent: A Design Framework for Music Learning and Beyond	2016
Medical Image Retrieval Using Bag of Meaningful Visual Words: Unsupervised Visual Vocabulary Pruning with PLSA	2013
Knowledge Audit for a Large Scale Government KM Strategy	2013
Event Digest: A Holistic View on Past Events	2016
End-user Mashup Programming: Through the Design Lens	2010
Agile Policy Development for Digital Government: An Exploratory Case Study	2013
Explicit Web Search Result Diversification	2012
Demystifying Magic: High-level Low-level Programming	2009
Optimal Incentive Timing Strategies for Product Marketing on Social Networks	2012
Selection Pressure and Takeover Time of Distributed Evolutionary Algorithms	2010
Removing Manually Generated Boilerplate from Electronic Texts: Experiments with Project Gutenberg e-Books	2007
Platform-dependent Code Generation for Embedded Real-time Software	2013
Knuckle Print Biometrics and Fusion Schemes – Overview, Challenges, and Solutions	2016
A Study of Querying Behaviour of Expert and Non-expert Users of Biomedical Search Systems	2014
Modeling Adaptation with a Tuple-based Coordination Language	2012
An Asymmetric Fingerprinting Code for Collusion-resistant Buyer-seller Watermarking	2013
Poor Video Streaming Performance Explained (and Fixed)	2016
Evolving Data Sets to Highlight the Performance Differences Between Machine Learning Classifiers	2012
An Analytical Model for Generalized Processor Sharing Scheduling with Heterogeneous Network Traffic	2007
Fibred Data Types	2013
Best Permutation: A Novel Network Selection Scheme in Heterogeneous Wireless Networks	2009
Who AM I?: Understanding High School Computer Science Teachers' Professional Identity	2012
A PSO-based Rank Aggregation Algorithm for Ranking Genes from Microarray Data	2013

Title	Year
Biomedical Knowledge Engineering Tools Based on Experimental Design: A Case Study Based on Neuroanatomical Tract-tracing Experiments	2009
Personalized and Contextualized Information in Self-management Systems for Chronically Ill Patients (PERISCOPE)	2010
Finding Social Points of Interest from Georeferenced and Oriented Online Photographs	2016
Five Perspectives on Computer Game History	2007
The Role of Music in the Lives of Homeless Young People in Seattle WA and Vancouver BC	2012
Evaluating Web Conferencing Tool Effectiveness	2011
Simulation Optimization for Industrial Scheduling Using Hybrid Genetic Representation	2008
Database Researchers: Plumbers or Thinkers?	2011
The Tower of Babel Meets Web 2.0: User-generated Content and Its Applications in a Multilingual Context	2010
On the Eeriness of Service Robots with Emotional Capabilities	2016
A Multiple Case Study of Small Free Software Businesses As Social Entrepreneurships	2015
Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection	2012
Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection	2012
Specifying Event-based Systems with a Counting Fluent Temporal Logic	2015
From Data Fusion to Knowledge Fusion	2014
Expectation Truncation and the Benefits of Preselection In Training Generative Models	2010
Reverse Annotation Based Retrieval from Large Document Image Collections	2010
Context-aware Agriculture Organizer	2012
Non Intrusive Load Monitoring: Systems, Metrics and Use Cases	2015
Click-based Hot Fixes for Underperforming Torso Queries	2016
Symbolic Analysis for Improving Simulation Coverage of Simulink/Stateflow Models	2008
YouPivot: Improving Recall with Contextual Search	2011
Towards Fine Grained Human Behaviour Simulation Models	2016
Creating Ontologies Using Ontology Mappings: Compatible and Incompatible Ontology Mappings	2013
Accelerating Linear System Solutions Using Randomization Techniques	2013
Toward a Better Understanding of Rule Initialisation and Deletion	2007
Ns-3 Based Framework for Simulating Communication Based Train Control (CBTC) Systems	2016
Partial Paradigm Hiding and Reusability in Hybrid Simulation Modeling Using the Frameworks health-DS and I7-AnyEnergy	2014
An Improved Corpus of Disease Mentions in PubMed Citations	2012
Investigating the Role of Top Management and Institutional Pressures in Cloud Computing Adoption	2016
Multi-Agent Task Assignment for Mobile Crowdsourcing Under Trajectory Uncertainties	2015
Modeling and Visualizing Uncertainty in Gene Expression Clusters Using Dirichlet Process Mixtures	2009
A Time and Space Complexity Analysis of Model Integration	2014
Reduced False Positives in PDZ Binding Prediction Using Sequence and Structural Descriptors	2012
Autonomous and Interdependent: Collaborative Privacy Management on Social Networking Sites	2016
A Study on the Management of a Discrete Event Simulation Project in a Manufacturing Company with PMBOK	2016
A Study on the Management of a Discrete Event Simulation Project in a Manufacturing Company with PMBOK	2016
Java Wildcards Meet Definition-site Variance	2012
Finding Top-k Similar Graphs in Graph Databases	2012
Detection of Outlier Residues for Improving Interface Prediction in Protein Heterocomplexes	2012
Fast Order-Preserving Encryption from Uniform Distribution Sampling	2015
Metrics for Phylogenetic Networks I: Generalizations of the Robinson-Foulds Metric	2009
Understanding the Limitations of Transmit Power Control for Indoor Wlans	2007
POSTER: Blind Separation of Benign and Malicious Events to Enable Accurate Malware Family Classification	2014
Brain-EE: Brain Enjoyment Evaluation Using Commercial EEG Headband	2016
Towards Estimating the Value of an Idea	2011
Comparison of Eye Movement Filters Used in HCI	2012
The Problem of Searching Interdisciplinary Learning Objects	2014
On Preserving Privacy in Content-oriented Networks	2011
Evidence-based Information Extraction for High Accuracy Citation and Author Name Identification	2007
Extending Two-variable Logic on Data Trees with Order on Data Values and Its Automata	2014
Blogger Perceptions on Digital Preservation	2007
Image Tagging with Social Assistance	2014
Variability Management in an Unaware Software Product Line Company: An Experience Report	2013
A Practical High Volume Software Product Line	2007
Facilitated Conceptual Modelling: Practical Issues and Reflections	2012
Reciprocity Attacks	2011
NoisySim: Exact Simulation of Stochastic Chemically Reacting Systems with Extrinsic Bounded Noises (WIP)	2013
Towards a Sustainable Crowdsourced Sound Heritage Archive by Public Participation: The Soundsslike Project	2016
Detecting and Understanding Students' Misconceptions Related to Algorithms and Data Structures	2012
Topological Relationship Query Processing for Complex Regions in Oracle Spatial	2012
An Automata Model for Trees with Ordered Data Values	2012
Empirical Analyses of the Factors Affecting Confirmation Bias and the Effects of Confirmation Bias on Software Developer/Tester Performance	2010
Lexical Normalization of Spanish Tweets	2016
Cluster Based Template Generation for Fast and Secure Fingerprint Matching	2016
Cloud Rhymer: Prototype Demo and Intervention Proposal	2013
The Complexity of Cross-boundary Information Sharing: An Organizational Perspective on Taiwan e-Government	2012
Online Portfolio Selection: A Survey	2014
Statistically-driven Alignment-based Multiword Expression Identification for Technical Domains	2009
Rapid Prototyping of Outdoor Games for Children in an Iterative Design Process	2013
A Unified Model for Joint Throughput-overhead Analysis of Mobile Ad Hoc Networks	2008
Knowledge Management and Organizational Culture in a Software Organization: A Case Study	2015
Touchstone: Exploratory Design of Experiments	2007
Indicator Selection for Daily Equity Trading with Recurrent Reinforcement Learning	2013
Towards a New Reading Experience via Semantic Fusion of Text and Music	2011

Title	Year
Boomerang: Resourceful Lenses for String Data	2008
Boomerang: Resourceful Lenses for String Data	2008
Identifying Linux Bug Fixing Patches	2012
BeUpright: Posture Correction Using Relational Norm Intervention	2016
Dynamic Structure for Web Graphs with Extended Functionalities	2016
Design of a Dynamic SLA Negotiation Protocol for Grids	2007
Communication Chains and Multitasking	2008
Measuring Wireless Network Success: An Analysis of a University in Ohio	2009
A Statistical Model of Cell-to-cell Variation in Li-ion Batteries for System-level Design	2013
Predictive Modeling of Drug Effects on Signaling Pathways in Diverse Cancer Cell Lines	2016
Data-driven Local Optima Network Characterization of QAPLIB Instances	2014
Political Use of Twitter: The Case of Metropolitan Mayor Candidates in 2014 Local Elections in Turkey	2014
Dynamic Execution of Secure Queries over Homomorphic Encrypted Databases	2016
Parameterised Verification of Autonomous Mobile-Agents in Static but Unknown Environments	2015
Automating Document Annotation Using Open Source Knowledge	2013
Are Computer Science Students Different Learners?	2010
Understanding the Process of Learning Touch-screen Mobile Applications	2013
On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation	2010
Understanding Career Choice of African American Men Majoring in Information Technology	2015
Evolution by Adapting Surrogates	2013
Exploring the Role of Mentoring in the IS Profession: A Cross-national Comparison	2012
Dynamic Rate Adaptation for Improved Throughput and Delay in Wireless Network Coded Broadcast	2014
Examining Vicinity Dynamics in Opportunistic Networks	2013
Relationships Between Reading, Tracing and Writing Skills in Introductory Programming	2008
Pairwise Structure Alignment Specifically Tuned for Surface Pockets and Interaction Interfaces	2015
Experiences with PDG-Based IFC	2010
Wireless Network MAC Layer Performance Evaluation with Full-Duplex Capable Nodes	2016
Reconstructing User's Attention on the Web Through Mouse Movements and Perception-Based Content Identification	2016
Probabilistic Biological Network Alignment	2013
Efficient Algorithms for Approximate Member Extraction Using Signature-based Inverted Lists	2009
Inverted Index Compression and Query Processing with Optimized Document Ordering	2009
Measuring the Impact of Third Place Attachment on the Adoption of a Place-based Community Technology	2009
Inquiry-based Learning in CS1	2007
Dwelling and Fleeting Encounters: Exploring Why People Use WeChat - A Mobile Instant Messenger	2015
Evaluation of E-Technology Acceptance: An Empirical Analysis	2012
Modeling Document Scores for Distributed Information Retrieval	2011
A Fuzzy Genetic Algorithm for the Dynamic Cell Formation Problem	2007
MEDUSA: MEtamorphic Malware Dynamic Analysis Usingsignature from API	2010
Research on Multi-supplier Performance Measurement Based on Genetic Ant Colony Algorithm	2009
Synthesising Liberal Normative Systems	2015
Weaving Executability into UML Class Models at PIM Level	2009
Directed and Heuristic Counterexample Generation for Probabilistic Model Checking: A Comparative Evaluation	2010
Designers Wanted: Participation and the User Experience in Open Source Software Development	2009
Human Tutorial Instruction in the Raw	2015
Audio-Based Multimedia Event Detection with DNNs and Sparse Sampling	2015
Toward a Framework for Ontology Modularity	2015
A Cosimulation Methodology for HW/SW Validation and Performance Estimation	2009
On the Effect of Black-hole Attack on Opportunistic Routing Protocols	2015
eLearning and Higher Education in Pakistan: What May Hamper It	2014
Information Horizons of College Students: Source Preferences and Source Referrals in Academic Contexts	2013
Formula-free Finite Abstractions for Linear Temporal Verification of Stochastic Hybrid Systems	2013
The Network in the Garden: An Empirical Analysis of Social Media in Rural Life	2008
Who's Viewed You?: The Impact of Feedback in a Mobile Location-sharing Application	2009
Some Usability Hypotheses for Verification	2015
Symbolic Documentation: Toward Fashion-related Sustainable Design	2012
Mobile Phone Adoption: Do Existing Models Adequately Capture the Actual Usage of Older Adults?	2010
Performance and Strain Levels of It Workers Engaged in Rapidly Changing Environments: A Person-job Fit Perspective	2010
Classifying the Humanities	2013
Implication in Information Retrieval Systems	2010
Using PageRank to Infer User Preferences	2012
Identifying Noun Product Features That Imply Opinions	2011
Organizational and Inter-organizational Precursors to Problematic Automation in Safety Critical Domains	2012
Simulation Implementation and Performance Analysis for Situational Awareness Data Dissemination in a Tactical Manet	2014
How Does Project Managers' Personality Matter?: Building the Linkage Between Project Managers' Personality and the Success of Software Development Projects	2009
User Model-based Metrics for Offline Query Suggestion Evaluation	2013
Bidding Strategy of a Microgrid in the Deregulated Electricity Market	2015
Generation of Process Using Multi-objective Genetic Algorithm	2013
Navigating Constraints: The Design Work of Professional Software Developers	2013
Provenance Query Evaluation: What's So Special About It?	2009
Bagging Gradient-boosted Trees for High Precision, Low Variance Ranking Models	2011
An Entropy Based Heuristic Model for Predicting Functional Sub-type Divisions of Protein Families	2009
What's the Date?: High Accuracy Interpretation of Weekday Names	2008
CLiSuite: Simplifying the Development of Cross-layer Adaptive Applications	2012
Can Recognising Multiword Expressions Improve Shallow Parsing?	2010
DREAM Project: Using Virtual Reality to Decrease Pain and Anxiety of Children with Burns During Treatments	2016
On the Number of Candidates in Opportunistic Routing for Multi-hop Wireless Networks	2013

Title	Year
Effective Automatic Parallelization of Stencil Computations	2007
Effective Automatic Parallelization of Stencil Computations	2007
Annotating Ecology: Looking to Biological Fieldwork for Mobile Spatial Annotation Workflows	2013
Undesired Relatives: Protection Mechanisms Against the Evil Twin Attack in IEEE 802.11	2014
Compatibility Between Trust and Non-Driving Related Tasks in UI Design for Highly and Fully Automated Driving	2016
On Quantile Regression in Reproducing Kernel Hilbert Spaces with the Data Sparsity Constraint	2016
On Improved CS-SS Image Watermark Detection over Radio Mobile Channel	2016
The Effect of Job Stress on Job Performance Amongst IT Professionals: The Moderating Role of Proactive Work Behaviours	2014
Minimizing Private Data Disclosures in the Smart Grid	2012
Teachers Collaborating with Wiki: The Impact of Professional Status, Language, and Age	2010
Quo Vadis, Clone Management?	2010
CAP3: Context-sensitive Abstract User Interface Specification	2011
Mapping Rules for Converting from ODL to XML Schemas	2010
Secure Random Number Generation in Wireless Sensor Networks	2011
Synaptic Channel Model Including Effects of Spike Width Variation	2015
Composition of Situational Interactive Spaces by End Users: A Case for Cultural Heritage	2012
The Ethics of Unaware Participation in Public Interventions	2016
Compact Full-text Indexing of Versioned Document Collections	2009
Combining Regression and Estimation by Analogy in a Semi-parametric Model for Software Cost Estimation	2008
Scientific Paper Summarization Using Citation Summary Networks	2008
Analysis of Ocean Clutter for Wide-band Radar Based on Real Data	2011
A Comprehensive Survey on Safe Reinforcement Learning	2015
Coordination in Emergency Response Management	2008
Symmetric Lenses	2011
Braingame: A Web-based Student Response System	2012
Towards an Understanding of Locality in Genetic Programming	2010
Teaching Computer Science in Secondary Education: A Technological Pedagogical Content Knowledge Perspective	2013
Simplified Partial Digest Problem: Enumerative and Dynamic Programming Algorithms	2007
Real-time Task Scheduling Using Extended Overloading Technique for Multiprocessor Systems	2007
Designing Visual Complexity for Dual-screen Media	2016
Analysis of Nonstationary Stochastic Simulations Using Classical Time-series Models	2009
Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI	2016
Smoothed Functional Algorithms for Stochastic Optimization Using q-Gaussian Distributions	2014
Let's Kick It: How to Stop Wasting the Bottom Third of Your Large Screen Display	2014
Supervisor Synthesis in Model-Based Automotive Systems Engineering	2014
Help-based Tutorials	2008
A Fully Automatic Hand Gesture Recognition System for Human-robot Interaction	2011
Bounds of Neglect Benevolence in Input Timing for Human Interaction with Robotic Swarms	2015
Extracting Information Networks from the Blogosphere	2012
Longitudinal Relationship Management for Instructional Technologists	2016
Volunteer Attraction and Retention in Open Source Communities	2014
An Analysis of Packet Sampling in the Frequency Domain	2009
A Stochastic Geometry Analysis of Multichannel Cognitive Radio Networks	2016
Learning Taxonomy Adaptation in Large-scale Classification	2016
An Experimental Investigation of Model-based Parameter Optimisation: SPO and Beyond	2009
The Tyrants, Heroes and Victims Narrative in Accessibility Tracks Papers of the Brazilian Symposium on Human Factors in Computer Systems	2016
Symbolic Computation of Differential Equivalences	2016
Solving Strategic Bargaining with Arbitrary One-sided Uncertainty	2011
The Meaning of IT Work	2013
Facilitating Collaboration Between COTS Stakeholders via Principles of Advanced ISD Methodologies: The Vendor Perspective	2015
On the Parameterized Complexity of Minimax Approval Voting	2015
Automated Design of Pin-constrained Digital Microfluidic Biochips Under Droplet-interference Constraints	2007
Improved Algorithms for Optimal Embeddings	2008
Convergence Results for Ant Routing Algorithms via Stochastic Approximation	2013
Information-Theoretic Lower Bounds on the Storage Cost of Shared Memory Emulation	2016
Designing Beyond Habit: Opening Space for Improved Recycling and Food Waste Behaviors Through Processes of Persuasion, Social Influence and Aversive Affect	2013
Scientists' Needs in Modelling Software Ecosystems	2015
The Design Space of Body Games: Technological, Physical, and Social Design	2013
Formal Modelling and Verification of the Common Hopping Multi-channel MAC Protocols	2010
QOOK: Enhancing Information Revisitation for Active Reading with a Paper Book	2013
Energy-efficient Fixed-priority Scheduling for Real-time Systems Based on Threshold Work-demand Analysis	2011
Walk This Way: Musically Guided Walking Experiences	2014
A New Semantic Functional Similarity over Gene Ontology	2015
Fairness-based Scheduling Algorithm for TDD Mode IEEE 802.16 Broadband Wireless Access Systems	2008
Equivalence and Minimization of Conjunctive Queries Under Combined Semantics	2012
Exploring Dependency for Query Privacy Protection in Location-based Services	2013
Improving Multibank Memory Access Parallelism with Lattice-Based Partitioning	2015
Modeling Network Traffic in Mobile Networks Implementing Offloading	2011
Churchman's Inquirers As Design Templates for Knowledge Management Systems	2009
Minimizing Test Suites in Software Product Lines Using Weight-based Genetic Algorithms	2013
Ground Soldier Materiel Task Related Encumbrance (MATRE) Model Creation Through the Use of Ontology Building and Analysis Tools	2009
Runtime Verification for LTL and TLTL	2011
Impact of Human Operators on Communication Network Dependability	2011
A Pattern Language for Service-based Platform Integration and Adaptation	2012

Title	Year
Self-tracking Cultures: Towards a Sociology of Personal Informatics	2014
Semantic Text Classification of Disease Reporting	2007
Examining the Viability of Mixed Framework for Evaluating Mobile Services Impact in Rural India	2010
Creativity in Video Game Design As Pedagogy	2009
Challenging the Supremacy of Traffic Matrices in Anomaly Detection	2007
A Demand-based e-Government Adoption Model (DeAM)	2016
Hardcore Casual: Game Culture Return(s) to Ravenhearst	2009
An Efficient Algorithm for Matching Protein Binding Sites for Protein Function Prediction	2011
Low Distortion Embeddings for Edit Distance	2007
The Drawing Assistant: Automated Drawing Guidance and Feedback from Photographs	2013
Emotion Detection State of the Art	2012
On the Terms Within- and Cross-Company in Software Effort Estimation	2016
Towards Exhaustive Protein Modification Event Extraction	2011
Stateful Aspects: The Case for Aspect-oriented Modeling	2007
Quantifying the Impact of Uncertainty in Human Actions on the Energy Performance of Educational Buildings	2016
Quantifying the Impact of Uncertainty in Human Actions on the Energy Performance of Educational Buildings	2016
Framework for Analysis of Online Course Design	2010
Residue Objects: A Challenge to Web Browser Security	2010
Compression in Wireless Sensor Networks: A Survey and Comparative Evaluation	2013
Impact of Software Resource Estimation Research on Practice: A Preliminary Report on Achievements, Synergies, and Challenges	2011
Indifference-zone Subset Selection Procedures: Using Sample Means to Improve Efficiency	2007
Action Research As a Model for Industry-academia Collaboration in the Software Engineering Context	2014
Modeling the Relationship Between Software Effort and Size Using Deming Regression	2010
Bayesian Synthesis for Knowledge Translation in Software Engineering: Method and Illustration	2016
A kit-based approach to preparing undergraduate students for research	2014
Data Sampling and Supervised Learning for HIV Literature Screening	2016
Exploring the architectural trade space of NASAs Space Communication and Navigation Program	2013
Development of maturity models: A systematic literature review	2012
MAnGve maturity model (M3): A proposal for a doctoral thesis	2014
Quality Measures for Documentation of Best Practices	2014
Sustainability in software engineering: A systematic literature review	2012
Variability in Software Systems—A Systematic Literature Review	2014
Empirical investigation of the challenges of the existing tools used in global software development projects	2015
The role of software process simulation modeling in software risk management: A systematic review	2009
Participatory design for the development of inclusive educational technologies: A systematic review	2016
Requirements elicitation approaches: A systematic review	2015
An Empirical Investigation of Systematic Reviews in Software Engineering	2011
A Framework for Classifying and Comparing Architecture-centric Software Evolution Research	2013
Common Agile Practices in Software Processes	2011
Process Model Engineering Lifecycle: Holistic Concept Proposal and Systematic Literature Review	2014
Costing of Cloud Computing Services: A Total Cost of Ownership Approach	2012
Notice of Violation of IEEE Publication Principles Systematically Identifying Genes and Pathways in Multiple Cancer Types Using HGD & PSO-SVM	2009
The Scalability-Efficiency/Maintainability-Portability Trade-Off in Simulation Software Engineering: Examples and a Preliminary Systematic Literature Review	2016
Supporting HIV literature screening with data sampling and supervised learning	2015
Evidence-based decision support for the clinical practice of acupuncture: Data mining approaches	2013
An empirical assessment of a systematic search process for systematic reviews	2011
Trends and perceptions of evidence-based software engineering research in Malaysia	2014
On the Use of Component-Based Principles and Practices for Architecting Cyber-Physical Systems	2016
Towards evidence-based ontology for supporting Systematic Literature Review	2012
Towards enhancing business process modeling formalisms of EKD with consistency consideration	2016
Hierarchical Fault Diagnosis and Health Monitoring in Satellites Formation Flight	2011
User-Centered Design and Agile Methods: A Systematic Review	2011
Sustainability evaluation framework for ambient intelligences mobile services	2013
Using Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews	2011
Techniques for the Identification of Crosscutting Concerns: A Systematic Literature Review	2016
Applications of ontologies in requirements engineering: a systematic review of the literature	2016
UML model refactoring: a systematic literature review	2015
Semi-automatic selection of primary studies in systematic literature reviews: is it reasonable?	2015
The Proactive and Reactive Digital Forensics Investigation Process: A Systematic Literature Review	2011
Awareness Support in Distributed Software Development: A Systematic Review and Mapping of the Literature	2013
Protocol-Driven Searches for Medical and Health-Sciences Systematic Reviews	2011
Features of IT Service Markets: A Systematic Literature Review	2016
A practice-driven systematic review of dependency analysis solutions	2011
A Systematic Review on the Impact of CK Metrics on the Functional Correctness of Object-Oriented Classes	2012
Understanding requirement prioritization artifacts: a systematic mapping study	2016
Survey on mining subjective data on the web	2012
Recovering from a decade: a systematic mapping of information retrieval approaches to software traceability	2014
Content-based Recommender Systems: State of the Art and Trends	2011
Mining unstructured content for recommender systems: an ensemble approach	2016
Context-Aware Recommender Systems	2015
Customer Feedback and Data Collection Techniques in Software R&D: A Literature Review	2015
The Effectiveness of Big Data in Health Care: A Systematic Review	2014
Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions	2015
The Impact of Knowledge Sharing Platforms in Distributed Requirements Engineering Scenarios: A Systematic Review	2014

Title	Year
Awareness Support in Global Software Development: A Systematic Review Based on the 3C Collaboration Model	2010
Hybreed: A software framework for developing context-aware hybrid recommender systems	2014
Mining textual requirements to assist architectural software design: a state of the art review	2012
Recommender Systems in Technology Enhanced Learning	2011
Data Network in Development of 3D Collaborative Virtual Environments: A Systematic Review	2014
How to Build a Recommendation System for Software Engineering	2015
Context-Aware Recommender Systems	2011
Webometrics benefitting from web mining? An investigation of methods and applications of two research fields	2014
Summarizing Software Artifacts: A Literature Review	2016
Layered evaluation of interactive adaptive systems: framework and formative methods	2010
MDE for BPM: A Systematic Review	2008
Systematic planning for digital preservation: evaluating potential strategies and building preservation plans	2009
Source Factors in Recommender System Credibility Evaluation	2015
Recommending people to people: the nature of reciprocal recommenders with a case study in online dating	2013
Requirements engineering education: a systematic mapping study	2015
Big data in manufacturing: a systematic mapping study	2015
A Systematic Review Measurement in Software Engineering: State-of-the-Art in Measures	2008
A data mining approach to guide students through the enrollment process based on academic performance	2011
Creating More Credible and Persuasive Recommender Systems: The Influence of Source Characteristics on Recommender System Evaluations	2011
Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication	2015
Bayesian networks for enhancement of requirements engineering: a literature review	2016
Reporting Empirical Research in Open Source Software: The State of Practice	2009
When does an easy task become hard? A systematic review of human task-evoked pupillary dynamics versus cognitive efforts	2016
A Personalized Method of Literature Recommendation Based on Brain Informatics Provenances	2015
Data Mining and Agent Technology: a fruitful symbiosis	2008
Multi-Criteria Recommender Systems	2015
Biomedical Text Mining: A Survey of Recent Progress	2012
A Self-organizing Sensing System for Structural Health Monitoring of Aerospace Vehicles	2008
Quality, productivity and economic benefits of software reuse: a review of industrial studies	2007
Introduction and Background	2013
Human Decision Making and Recommender Systems	2015
Pragmatic Interoperability: A Systematic Review of Published Definitions	2010
A people-to-people matching system using graph mining techniques	2014
The social bookmark and publication management system bibsonomy	2010
Formation and Fragmentation	2011
Unveiling the intellectual origins of Social Media-based innovation: insights from a bibliometric approach	2016
A Literature Review on Cloud Computing Adoption Issues in Enterprises	2014
Ontology-Based Model for Mining User's Emotions on the Wisdom Web	2016
State-of-the-art in biomedical literature retrieval for clinical cases: a survey of the TREC 2014 CDS track	2016
Evaluating Recommendation Systems	2011
A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law	2012
Evaluating Recommender Systems	2015
The evaluation of a social adaptive website for cultural events	2013
A Quarter of Century in Artificial Intelligence and Law: Projects, Personal Trajectories, a Subjective Perspective	2014
Software Engineering – Processes and Tools	2009
Recommender Systems	2010
Evaluation of Engineering Approaches in the Secure Software Development Life Cycle	2014
Finding best evidence for evidence-based best practice recommendations in health care: the initial decision support system design	2011
Biomedical Text Mining: State-of-the-Art, Open Problems and Future Challenges	2014
Mining system logs to learn error predictors: a case study of a telemetry system	2015
Rough set-based approaches for discretization: a compact review	2015
Ecosystems and Open Innovation for Embedded Systems: A Systematic Mapping Study	2015
An Investigation into Agile Methods in Embedded Systems Development	2012
Context-Aware Recommender Systems Influenced by the Users' Health-Related Data	2013
Towards a Holistic Approach for Sustainable Partner Selection in the Electrics and Electronics Industry	2011
Research Themes in the Case-Based Reasoning in Health Sciences Core Literature	2012
Understanding big data themes from scientific biomedical literature through topic modeling	2016
Quality of Business Process Models	2016
A review of data mining using big data in health informatics	2014
Integration of decision support systems to improve decision support performance	2010
Context Modeling, Representation, and Reasoning: An Ontological and Hybrid Approach	2015
General Context and Theories	2013
Multimedia Data Mining	2010
CyberGate Visualization	2012
Empirical studies in reverse engineering: state of the art and future trends	2007
A survey on enhanced subspace clustering	2013
Use of the Global Trigger Tool in patient safety improvement efforts: Nordic experiences	2015
The Practice Areas	2011
Towards a decision-making structure for selecting a research design in empirical software engineering	2015
Data Mining and Information Systems: Quo Vadis?	2010
Users' Information Search Behavior in a Professional Search Environment:	2014
Quality Models	2013
Language use reflects scientific methodology: A corpus-based study of peer-reviewed journal articles	2008
Towards an effective automatic query expansion process using an association rule mining approach	2012

Title	Year
Data Mining and Data Fusion for Enhanced Decision Support	2008
Context-based literature digital collection search	2009
Automated Log Audits for Privacy Compliance Validation: A Literature Survey	2016
More Than Just Black and White: A Case for Grey Literature References in Scientific Paper Information Retrieval Systems	2015
Large-scale information retrieval in software engineering - an experience report from industrial application	2016
A Survey of Commonsense Knowledge Acquisition	2013
Sourcerer: mining and searching internet-scale software repositories	2009
Preferences in artificial intelligence	2016
Unpacking the Artifact Knowledge: Secondary Data Analysis in Design Science Research	2015
Mining Videos from the Web for Electronic Textbooks	2014
An Overview of Social Tagging and Applications	2011
Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation	2008
Madonna: a large-scale evaluation on estimating similarities between music artists and between movies from microblogs	2012
Do Persuasive Technologies Persuade? - A Review of Empirical Studies	2014
Multimedia Information Networks in Social Media	2011
Semantics discovery in social tagging systems: A review	2016
An industry experiment on the effects of test-driven development on external quality and productivity	2016
Simulating the Past for Understanding the Present. A Critical Review	2016
Information Retrieval in Literature-Based Discovery	2008
Thinking Out of the Box: Discovering the Relevance of External Context to Business Processes	2013
Putting Value into E-valu-ation	2008
Fuzzy Methods for Data Mining and Machine Learning: State of the Art and Prospects	2008
An Ontology-Learning Knowledge Support System to Keep e-Organization's Knowledge Up-to-Date: A University Case Study	2011
Improving the Applicability of Environmental Scanning Systems: State of the Art and Future Research	2011
A Survey of Graph Mining Techniques for Biological Datasets	2010
Understanding the role of latent feature diversification on choice difficulty and satisfaction	2016
Managing Requirements Knowledge: Conclusion and Outlook	2013
Research on Quality of Transaction Standards: The Maturity of a Research Topic	2010
A Procedure of How to Conduct Research in Transparent Mobile Recommendations	2008
Decision and stress: cognition and e-accessibility in the information workplace	2007
Mapping and Usage of Know-How Contributions	2015
Machine Translation from Text	2011
Text Mining Scientific Papers: A Survey on FCA-Based Information Retrieval Research	2012
Visual requirements analytics: a framework and case study	2014
Investigating task performance of probabilistic topic models: an empirical study of PLSA and LDA	2011
Modelling and Mining of Networked Information Spaces	2008
A Case Study to Enable and Monitor Real IT Companies Migrating from Waterfall to Agile	2014
Semantic tagging and linking of software engineering social content	2016
Response-Based Segmentation Using Finite Mixture Partial Least Squares	2010
Rationale Management and Traceability in Detailed Discussion	2013
From World-Wide-Web Mining to Worldwide Webmining: Understanding People's Diversity for Effective Knowledge Discovery	2007
Detecting New Evidence for Evidence-Based Guidelines Using a Semantic Distance Method	2015
Profiling relational data: a survey	2015
Architecture and Quality in Data Warehouses	2013
Identifying Seekers and Suppliers in Social Media Communities to Support Crisis Coordination	2014
A review of information resources on nanoscience, nanotechnology, and nanomaterials	2015
Exploring Extremism and Terrorism on the Web: The Dark Web Project	2007
The Status Quo of Ontology Learning from Unstructured Knowledge Sources for Knowledge Management	2014
A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements	2011
Abstraction in Machine Learning	2013
An emotional polarity analysis of consumers' airline service tweets	2013
The CACHE Study: Group Effects in Computer-supported Collaborative Analysis	2008
Filter Approach Feature Selection Methods to Support Multi-label Learning Based on ReliefF and Information Gain	2012
Information and Semantics in Databases and on the Web	2009
The Science and Art of Conceptual Modelling	2012
Evaluating Business Intelligence Gathering Techniques for Horizon Scanning Applications	2013
On the impact of software evolution on software clustering	2013
Towards a Procedure Model for Developing Anaphora Processing Applications	2016
A Review of the Theory of Planned Behaviour in the Context of Information Security Policy Compliance	2013
Anatomy of Knowledge Bases Used in Design Science Research	2012
Actionable Knowledge As A Service (AKAAS): Leveraging big data analytics in cloud computing environments	2015
Scenario-Based Assessment of Process Pattern Languages	2009
Data Classification	2015
High Quality Information Retrieval for Improving the Conduct and Management of Research and Development	2010
A methodology for the selection of requirements engineering techniques	2008
Safeguards	2010
Modelling Languages: Perspectives and Abstraction Mechanisms	2012
Robust personalizable spam filtering via local and global discrimination modeling	2013
Towards Machine Learning on the Semantic Web	2008
Can predicate-argument structures be used for contextual opinion retrieval from blogs?	2013
A trace-based approach to identifying users' engagement and qualifying their engaged-behaviours in interactive systems: application to a social game	2014
The Trouble with 'Tacit Knowledge'	2012
Interacting with Context	2009
Patterns of Work: A Pragmatic Approach	2015
Subspace and projected clustering: experimental evaluation and analysis	2009

Title	Year
Pattern Clustering Using a Swarm Intelligence Approach	2010
A survey of learning-based techniques of email spam filtering	2008
Knowledge Discovery and Expert Knowledge for Creating a Chart-Model of a Biological Network - Introductory to Research in Chronic Diseases and Comorbidity	2011
Omissions in Managing Knowledge in Innovation Processes or How to Handle Knowledge, Humans and Tasks: A Semio-Cognitive Approach	2007
Literature-related discovery: common factors for Parkinson's Disease and Crohn's Disease	2014
Beyond Bitcoin Enabling Smart Government Using Blockchain Technology	2016
Computational Systems Biology Approaches for Deciphering Traditional Chinese Medicine	2013
Impact of Text Mining Application on Financial Footnotes Analysis	2015
A Unified View of Data-Intensive Flows in Business Intelligence Systems: A Survey	2016
Visualisation of Integrated Patient-Centric Data as Pathways: Enhancing Electronic Medical Records in Clinical Practice	2016
The Service Science Foundation for Cloud Computing	2016
Bus Bunching Detection by Mining Sequences of Headway Deviations	2012
Applications of Semantic Web Methodologies and Techniques to Biology and Bioinformatics	2008
Using Social Media to Empower Learning Resources Evaluation and Recommendation Across Boundaries	2011
Cognitive Process	2014
Multimedia, Broadcasting, and eCulture	2011
Human Involvement in Designing an Information Quality Assessment Technique – Demonstrated in a Healthcare Setting –	2012
SMILI: a Framework for Interfaces to Learning Data in Open Learner Models, Learning Analytics and Related Fields	2016
Computational Intelligence in Multimedia Processing: Foundation and Trends	2008
Computational intelligence in biomedical imaging: multidimensional analysis of spatio-temporal patterns	2011
Creating a model of the dynamics of socio-technical groups	2013
Cluster Analysis	2015
Unraveling Challenges in Collaborative Design: A Literature Study	2009
Collected Papers	2009
Development and Implementation of a Safety Management System in a Lean Airline	2015
New Ways of Working: A Proposed Framework and Literature Review	2011
Foraging	2012
From global to local and viceversa: uses of associative rule learning for classification in imprecise environments	2012
User Methods and Approaches to Design Cognitive Systems	2016
Massive Scientific Paper Mining: Modeling, Design and Implementation	2013
Challenges for Migrating to the Service Cloud Paradigm: An Agile Perspective	2013
Assessing architectural evolution: a case study	2011
A practical approach to assess risk in aviation domains for safety management systems	2015
Future Patent Search	2011
US Domestic Extremist Groups	2012
Automatic image semantic interpretation using social action and tagging data	2011
Towards Crowd-Based Requirements Engineering A Research Preview	2015
Intelligent Knowledge Management	2011
Chapter 8. Experiments	2010
Managing Legal Texts in Requirements Engineering	2009
Human-Centered Aspects	2007
Swarm Intelligence Algorithms for Data Clustering	2008
Exploration of Two-Objective Scenarios on Supervised Evolutionary Feature Selection: A Survey and a Case Study (Application to Music Categorisation)	2015
A signed-distance-based approach to importance assessment and multi-criteria group decision analysis based on interval type-2 fuzzy set	2013
Designing Academic Writing Analytics for Civil Law Student Self-Assessment	2016
Sentiment Analysis: What's Your Opinion?	2014
Ecosystem Services Assessment Using InVEST as a Tool to Support Decision Making Process: Critical Issues and Opportunities	2015
A Self-Service MSS Design from a New-Generation Manager Perspective	2014
Detecting, classifying, and tracing non-functional software requirements	2016
Enterprise Social Networks: Status Quo of Current Research and Future Research Directions	2016
Do Teams Achieve Usability Goals? Evaluating Goal Achievement with Usability Goals Setting Tool	2011
Evaluating ReliefF-Based Multi-Label Feature Selection Algorithm	2014
Applications and Experiences	2009
Topics in Software Industry Transformation Research: A Topic Analysis of Major IS Conferences	2012
What are Information Security Ontologies Useful for?	2015
Information Quality Framework for Verifiable Intelligence Products	2010
A survey on data stream clustering and classification	2015
Ontology Alignment Evaluation Initiative: Six Years of Experience	2011
Boosting support vector machines for imbalanced data sets	2010
Are delayed issues harder to resolve? Revisiting cost-to-fix of defects throughout the lifecycle	2016
Metaheuristic Pattern Clustering – An Overview	2009
Towards Construction of Situational Methods for Service Identification	2011
Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems	2010
Contribution of Geomatics Engineering and VGI Within the Landslide Risk Assessment Procedures	2015
Visual Analytics	2013
Empirical Insights on a Value of Ontology Quality in Ontology-Driven Web Search	2008
Spam filtering: how the dimensionality reduction affects the accuracy of Naive Bayes classifiers	2011
Social media and political communication: a social media analytics framework	2013
Analyzing Clusters and Constellations from Untwisting Shortened Links on Twitter Using Conceptual Graphs	2013
Architecture Enforcement Concerns and Activities - An Expert Study	2016
An empirical study of notifications' importance for online social network users	2015
Mining Anti-coagulant Drug-Drug Interactions from Electronic Health Records Using Linked Data	2013
Semantic topic models for source code analysis	2016

Title	Year
ESA Knowledge Management Agenda	2015
Participatory Design of Public Sector Services	2010
How Enterprise Architecture Formative Critical Success Facets Might Affect Enterprise Architecture Success: A Literature Analysis	2014
Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization	2009
A Compliance Management Ontology: Developing Shared Understanding through Models	2012
Functional, Control and Data Flow, and Mutation Testing: Theory and Practice	2010
'Technology Mining' bibliometrics analysis: applying network analysis and cluster analysis	2015
Meso-level retrieval: IR-bibliometrics interplay and hybrid citation-words methods in scientific fields delineation	2015
Assessing the Usability of Video Browsing and Summarization Techniques	2009
Information Extraction from Text	2012
Incremental data-driven learning of a novelty detection model for one-class classification with application to high-dimensional noisy data	2009
Computational Theory Perception (CTP), Rough-Fuzzy Uncertainty Analysis and Mining in Bioinformatics and Web Intelligence: A Unified Framework	2010
Techno-Organizational Diversity, Network Topologies and the Manageability of Banks	2010
2-Tuple Linguistic Decision Based Applications	2015
Using Models at Runtime to Address Assurance for Self-Adaptive Systems	2014
Mining Outliers with Ensemble of Heterogeneous Detectors on Random Subspaces	2010
Writing Clinical Practice Guidelines in Controlled Natural Language	2010
Social networks on the Internet	2013
Comparing Software Architecture Descriptions and Raw Source-Code: A Statistical Analysis of Maintainability Metrics	2013
Personalized Diversity Search Based on User's Social Relationships	2012
Introduction to Information Quality	2016
Evaluation of digital libraries	2007
Business Process Design: Towards Service- Based Green Information Systems	2008
Interactive Multiobjective Optimization from a Learning Perspective	2008
Formal specification of non-functional properties of component-based software systems	2010
Detecting the intellectual structure of library and information science based on formal concept analysis	2015
Translating Theory into Methodological Practice	2016
Requirements-Driven Social Adaptation: Expert Survey	2014
Towards Ontology-Based Information Systems and Performance Management for Collaborative Enterprises	2015
Sensor Deployments for Home and Community Settings	2013
Modeling score distributions in information retrieval	2011
Granular Computing, Philosophical Foundation for	2012
A Visual Analysis Approach for Community Detection of Multi-Context Mobile Social Networks	2013
Code clones and developer behavior: results of two surveys of the clone research community	2016
Improving Classifier Performance by Knowledge-Driven Data Preparation	2012
Extracting Novel Features for E-Commerce Page Quality Classification	2013
Knowledge-oriented software engineering process in a multi-cultural context	2010
Extracting common emotions from blogs based on fine-grained sentiment clustering	2011
Impact of preprocessing on medical data classification	2016
Improving Business Process Model after Reverse Engineering	2013
Interpreting communities based on the evolution of a dynamic attributed network	2015
A Pseudo-Relevance Feedback Based Method to Find Comprehensive Web Documents	2013
A study of the effect of different types of noise on the precision of supervised learning techniques	2010
Health Care Performance Monitoring Using an Inpatient Reference Process Model and Clinical KPIs	2012
Collaborative Computational Intelligence in Economics	2009
On artificial immune systems and swarm intelligence	2010
Towards the Development of a Framework for Encouraging the Learning of SPICE Model by Using Knowledge Graphs	2015
Digital Drugs	2014
Training the Body: The Potential of AIED to Support Personalized Motor Skills Learning	2016
Automatic detection of learner's affect from conversational cues	2008
Improving daily deals recommendation using explore-then-exploit strategies	2015
Uncertainty in Spatial Trajectories	2011
Evidence-Based Clinical Guidelines in SemanticCT	2014
A Critical Analysis of EU-Funded eParticipation	2012
Introduction to High-Dimensionality	2015
Dark scenarios	2010
The Added Value of Argumentation	2013
Deception detection: dependable or defective?	2014
Strengthening Web Based Learning through Software Quality Analysis	2009
A Short Survey on Online and Offline Methods for Search Quality Evaluation	2016
Coopetitive Data Warehouse: A Case Study	2013
On the role of semantics in automated requirements tracing	2015
From misuse cases to mal-activity diagrams: bridging the gap between functional security analysis and design	2014
On Intertext in Chemotherapy: an Ethnography of Text in Medical Practice	2016
Vision and Evolution: State of the Art	2016
Algorithm Selection for Combinatorial Search Problems: A Survey	2016
Dark Web Attribute System	2012
Machine Learning Methods in Automatic Image Annotation	2010
Technology Roadmap Development for Big Data Healthcare Applications	2015
On eliciting requirements from end-users in the ICT4D domain	2011
Emotions in autonomous agents: comparative analysis of mechanisms and functions	2012
A Study in Empirical and 'Casuistic' Analysis of Ontology Mapping Results	2007
Introduction and Challenges of Environment Architectures for Collective Intelligence Systems	2015
Implications of an ethic of privacy for human-centred systems engineering	2008

Title	Year
Identifying Illegal Cartel Activities from Open Sources	2016
The Virtuous Circle of the Quantified Self: A Human Computational Approach to Improved Health Outcomes	2013
Data Driven Governments: Creating Value Through Open Government Data	2016
Natvised, Playfully Aetiologised Literary Zoonyms, III: Abdim's Stork. Substituted Eponym, Dense Cultural Rewiring, Ethics	2014
On the Complexity of Requirements Engineering for Decision-Support Systems: The CID Case Study	2015
A vision towards Scientific Communication Infrastructures	2013
Combining formal concept analysis and translation to assign frames and semantic role sets to French verbs	2014
Graffiti: graph-based classification in heterogeneous networks	2012
The MACS Project: An Approach to Affordance-Inspired Robot Control	2008
A Facet-Based Methodology for the Construction of a Large-Scale Geospatial Ontology	2012
A High-Throughput Bioinformatics Platform for Mass Spectrometry-Based Proteomics	2007
Introduction	2010
Experiences and Lessons Learned with the Development of a Source Code Search Engine	2013
Is Knowledge Power? The Role of Knowledge in Automated Requirements Elicitation	2013
Domain Specific Search	2014
The Nature and a Process for Development of Enterprise Architecture Principles	2015
Healthcare Knowledge Management: The Art of the Possible	2008
The Name and Nature of Software Engineering	2008
Trust in Computing	2013
The Generation of Experimental Data for Computational Testing in Optimization	2010
Hybrid genetic algorithm for dual selection	2008
Configuring latent Dirichlet allocation based feature location	2014
From Sensor Readings to Predictions: On the Process of Developing Practical Soft Sensors	2014
Web Engineering at the Frontier of the Web 2.0: Design Patterns for Online 3D Shared Spaces	2009
Exploring Training Issues in Healthcare: Towards Identifying Barriers to Increase Electronic Medical Records Adoption by Healthcare Professionals	2011
Analysing Incomplete Consumer Web Data Using the Classification and Ranking Belief Simplex (Probabilistic Reasoning and Evolutionary Computation)	2010
Development of an Augmented Feedback Application to Support Motor Learning after Stroke: Requirement Analysis	2011
A sparse kernel relevance model for automatic image annotation	2014
Lean Software Development in Action	2014
Controlled experiments on the web: survey and practical guide	2009
Group Types in Social Media	2015
Simulating Simple and Fallible Relevance Feedback	2011
Assessing Medical Treatment Compliance Based on Formal Process Modeling	2011
Context Dependency Management in Ontology Engineering: A Formal Approach	2007
IT Artifact and IS Development	2013
TCO Evaluation in Physical Asset Management: Benefits and Limitations for Industrial Adoption	2014
Three Cases of Feature-Based Variability Modeling in Industry	2014
Best practices for learning video concept detectors from social media examples	2015
Field Experiments in HCI: Promises and Challenges	2009
Empirical Paradigm – The Role of Experiments	2007
Ontologies for Cultural Heritage	2009
Identifying and addressing problems in object-oriented framework reuse	2007
Modeling Control Objectives for Business Process Compliance	2007
Simulation	2014
The SmarterContext Ontology and Its Application to the Smart Internet: A Smarter Commerce Case Study	2013
Combining Traditional Cyber Security Audit Data with Psychosocial Data: Towards Predictive Modeling for Insider Threat Mitigation	2010
Cognition or Affect? - Exploring Information Processing on Facebook	2011
Mapping Dark Web Geolocation	2008
The Importance of Architectural Knowledge in Integrating Open Source Software	2011
Characterization of Cyber-Foraging Usage Contexts	2015
A Process Model of KMS Adoption and Diffusion in Organization: An Exploratory Study	2011
Advanced CBR Elements	2013
Naming the pain in requirements engineering	2016
Unintended Consequences of Class-Based Ontological Commitment	2011
Web Search Tailored Ontology Evaluation Framework	2007
Genetic operators for combinatorial optimization in TSP and microarray gene ordering	2007
7 Requirements Modeling for Embedded Realtime Systems	2010
When Cutting Edge Technology Meets Clinical Practice: Ethical Dimensions of e-Health	2014
Personalized and Adaptive Serious Games	2016
On Significance of Ontology Quality in Ontology-Driven Web Search	2008
The differentiation of the strategic profile of higher education institutions. New positioning indicators based on microdata	2008
Most Relevant Explanation: computational complexity and approximation methods	2011
Spatial Game Analytics	2013
Developer initiation and social interactions in OSS: A case study of the Apache Software Foundation	2015
Context and Fusion: Definitions, Terminology	2016
Apogee: Application Ontology Generation with Size Optimization	2015
Model-based engineering for change-tolerant systems	2007
New forms of interaction and knowledge sharing on Web 2.0	2009
Using 'core documents' for detecting and labelling new emerging topics	2012
Evaluation of visual video summaries: user-supplied constructs and descriptions	2011
Topic-based ranking in Folksonomy via probabilistic model	2011
Interaction and User Interfaces	2012
Content-Based Audio Retrieval	2015
Research profiling for 'standardization and innovation'	2011

Title	Year
Multidimensional networks: foundations of structural analysis	2013
Cognitive computing and proposed approaches to conceptual organization of case law knowledge bases: a proposed model for information preparation, indexing, and analysis	2016
Ambient Intelligence and Smart Environments: A State of the Art	2010
Towards Identifying the Business Value of Big Data in a Digital Business Ecosystem: A Case Study from the Financial Services Industry	2016
Harvesting and analysis of weak signals for detecting lone wolf terrorists	2013
Evaluating Citation Functions in CiTO: Cognitive Issues	2014
Assessment of Online PPGIS Study Cases in Urban Planning	2012
Chapter 3. Selected Design Issues	2010
Early warning analysis for social diffusion events	2012
The Organization of Knowledge	2015
Extracting and quantifying eponyms in full-text articles	2014
Retrieval Effectiveness of Cross Language Information Retrieval Search Engines	2011
Research on Chinese negation and speculation: corpus annotation and identification	2016
Lest we forget	2011
Value Added by Interoperable Information Systems in Spread Production Networks	2010
Learning to suggest questions in social media	2015
Topic-based heterogeneous rank	2015
Deliberation dialogues for reasoning about safety critical actions	2012
Can Shared-Neighbor Distances Defeat the Curse of Dimensionality?	2010
Faster issue resolution with higher technical quality of software	2012
Drafting a Composite Indicator of Validity for Regulatory Models and Legal Systems	2014
Managing dimensionality in data privacy anonymization	2016
Models for Network Graphs	2009
Reconciling Theories with Design Choices in Design Science Research	2013
The Bologna reform at the department of library and information science and book studies, university of Ljubljana	2009
Integrating a Bottom–Up and Top–Down Methodology for Building Semantic Resources for the Multilingual Legal Domain	2010
Integrating Medical Scientific Knowledge with the Semantically Quantified Self	2016
Collaborative video searching on a tabletop	2007
Understanding Understandability of Conceptual Models – What Are We Actually Talking about?	2012
KIPO: the knowledge-intensive process ontology	2015
Pattern-Based Specification of Crowdsourcing Applications	2014
The Three Steps of Clustering in the Post-Genomic Era: A Synopsis	2011
Social Knowledge Management in Practice: A Case Study	2009
State of Online Privacy: A Technical Perspective	2012
Supporting Business Processes for Collaborative Alliances of Software Service Providers	2015
Games for Health	2016
The THESEUS Use Cases	2014
Dark Web Research Overview	2012
Market Engineering: A Research Agenda	2008
Ontological Knowledge Maintenance Methodology	2008
Discovering Low Overlapping Biclusters in Gene Expression Data Through Generic Association Rules	2015
The Digital Coral: Infrastructuring Environmental Monitoring	2015
Evaluation of Hospital Portals Using Knowledge Management Mechanisms	2007
Evolution styles: foundations and models for software architecture evolution	2014
Developing a prototype information system framework to handle pedagogical knowledge in a secondary school modern foreign languages department	2008
Predicting Customer Loyalty Labels in a Large Retail Database: A Case Study in Chile	2010
Introduction	2010
Combining Textual and Visual Information for Semantic Labeling of Images and Videos	2008
Discourse Coherence: Lexical Chain, Complex Network and Semantic Field	2013
Task Analysis for Behavioral Factors Evaluation in Work System Design	2011
Modeling the Web of Data (Introductory Overview)	2011
A visual digital library approach for time-oriented scientific primary data	2011
Video personalization in heterogeneous and resource-constrained environments	2011
Deciding to Adopt Requirements Traceability in Practice	2007
FURIA: an algorithm for unordered fuzzy rule induction	2009
Measuring KM Success and KM Service Quality with KnowMetrix – First Experiences from a Case Study in a Software Company	2009
Analysis of attribute weighting heuristics for analogy-based software effort estimation method AQUA+	2008
Eradicating root causes of aviation maintenance errors: introducing the AMMP	2014
Informative Value of Individual and Relational Data Compared Through Business-Oriented Community Detection	2013
Stacking Label Features for Learning Multilabel Rules	2014
Breaking the Recursivity: Towards a Model to Analyse Expert Finders	2015
Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method	2014
Cyber-physical-social-thinking space based science and technology framework for the Internet of Things	2015
Neurocognitive Approach to Clustering of PubMed Query Results	2009
From knowledge-based to data-driven fuzzy modeling	2015
Teaching Information Visualization	2008
Realizing the Innovation Potentials from Open Data: Stakeholders' Perspectives on the Desired Affordances of Open Data Environment	2016
Native and Multiple Targeted Mobile Applications	2015
Towards Management of the Data and Knowledge Needed for Port Integration: An Initial Ontology	2014
Two Step graph-based semi-supervised Learning for Online Auction Fraud Detection	2015
Towards Bottom-Up Decision Making and Collaborative Knowledge Generation in Urban Infrastructure Projects Through Online Social Media	2015
The pursuit of academic excellence and business engagement: is it irreconcilable?	2013

Title	Year
Formal Methods as a Link between Software Code and Legal Rules	2011
Finding Repeated Patterns in Music: State of Knowledge, Challenges, Perspectives	2014
A Cognitive Support Framework for Ontology Mapping	2007
Applications of Multilingual Information Access	2012
Exploring mindlessness as an explanation for the media equation: a study of stereotyping in computer tutorials	2009
SHIRAZ: an automated histology image annotation system for zebrafish phenomics	2011
The GRIFFIN Collaborative Virtual Community for Architectural Knowledge Management	2010
Designing Consumer Health Information Systems: What Do User-Generated Questions Tell Us?	2011
Similarity measures for OLAP sessions	2014
Tailoring Interface for Spanish Language: A Case Study with CHICA System	2009
Syntactic sentence compression in the biomedical domain: facilitating access to related articles	2007
Creating Realistic Topics for Image Retrieval Evaluation	2010
Geometry and Orientation of the Single Image	2016
Networked Intelligent Retrieval System Based on Semantic	2013
Learning-Based Interactive Retrieval in Large-Scale Multimedia Collections	2013
From Water Music to 'Underwater Music': Multimedia Soundtrack Retrieval with Social Mass Media Resources	2016
Introduction	2011
Human Computation and Crowdsearching	2013
An Empirically Informed Taxonomy for the Maker Movement	2016
Performance Effects of Positive and Negative Affective States in a Collaborative Information Seeking Task	2014
Decisions and Knowledge	2008
Real-World Applications of Multiobjective Optimization	2008
Whistle for music: using melody transcription and approximate string matching for content-based query over a MIDI database	2007
Improving Medical Information Retrieval with PICO Element Detection	2010
Evolutionary Algorithm Parameters and Methods to Tune Them	2012
Multi-Agent Fuzzy-Based Control Architecture for Autonomous Mobile Manipulators: Traditional Approaches and Multi-Agent Fuzzy-Based Approaches	2013
Measuring gene similarity by means of the classification distance	2011
Staff	2014
Complex Values-Based Approach for Multidimensional Evaluation of Landscape	2014
Introduction: What Landmarks Are, and Why They Are Important	2014
Technology and Care for Patients with Chronic Conditions: The Chronic Care Model as a Framework for the Integration of ICT	2012
Inferring Phylogenetic Trees Using a Multiobjective Artificial Bee Colony Algorithm	2012
Investigating Translator-Information Interaction: A Case Study on the Use of the Prototype Biconcordancer Tool Integrated in CASMACAT	2016
Adapting Progress Feedback and Emotional Support to Learner Personality	2016
Interactive Decision Aids	2012
ICT Services for open and citizen science	2015
Case-Based Decision Making	2007
Learning Theory and Algorithmic Quality Characteristics	2016
Infosel++: Information Based Feature Selection C++ Library	2010
Multimodal biomedical image retrieval using hierarchical classification and modality fusion	2013
Knowledge Intensive Business Processes: A Process-Technology Fit Perspective	2010
A Facet-Based Methodology for Geo-Spatial Modeling	2011
Internet practice and professional networks in Chilean science: Dependency or progress?	2009
Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting	2014
IaaS Adoption Determinants in Enterprises	2010
End Users Programming Smart Homes – A Case Study on Scenario Programming	2013
Quantum Information, Games and Computation	2010
A core ontology on events for representing occurrences in the real world	2012
Senior citizens and the ethics of e-inclusion	2009
Advances in computational facial attractiveness methods	2016
Beyond Concept Learning	2012
Sequential and Parallel Variable Neighborhood Search Algorithms for Job Shop Scheduling	2008
Recent progress and challenges in exploiting graphics processors in computational fluid dynamics	2014
From enterprise architecture to business models and back	2014
Modelling collective decision making in groups and crowds: Integrating social contagion and interacting emotions, beliefs and intentions	2013
A Taxonomy of Asymmetric Requirements Aspects	2007
Rosso Tiziano: A System for User-Centered Exploration and Discovery in Large Image Information Bases	2008
Analyzing the Accuracy of Calculations When Scoping Product Configuration Projects	2012
Optimization Problems	2011
An ontology-based evidential framework for video indexing using high-level multimodal fusion	2014
Ontology Engineering – The DOGMA Approach	2009
A Pool of Topics: Interactive Relational Topic Visualization for Information Discovery	2010
Introduction	2014
Analyzing Medical Processes	2011
An Organizational Model for Digital Library Evaluation	2011
Performance evaluation of incremental training method for face recognition using PCA	2007
Achieving Maturity: The State of Practice in Ontology Engineering in 2009	2009
Little Design Up-Front: A Design Science Approach to Integrating Usability into Agile Requirements Engineering	2009
Making Computer Learning Easier for Older Adults: A Community Study of Tuition Practices	2010
What Methodology Attributes Are Critical for Potential Users? Understanding the Effect of Human Needs	2011
An Open Data Approach for Clinical Appropriateness	2015
Using Mobile Devices in Supervision of Graduate Research in Distance Education: A Personal Journey	2015
Social Signal Processing: The Research Agenda	2011
Innovative Integrated Architecture for Educational Games: Challenges and Merits	2011

Title	Year
Designing Enterprise System Information Architecture Using Task Data	2015
A web-based platform for biosignal visualization and annotation	2014
Computational Intelligence in Economics and Finance: Shifting the Research Frontier	2007
VML* – A Family of Languages for Variability Management in Software Product Lines	2010
Assessing the Scholarly Impact of ImageCLEF	2011
Discussion on the Challenges and Opportunities of Cloud Forensics	2012
Methodologies and tools for audio digital archives	2009
Towards a Service-Oriented Methodology: Business-Driven Guidelines for Service Identification	2007
Knowledge Audit on Special Children Communities	2009
"Machine Beauty" – Should It Inspire eHealth Designers?	2014
Analyzing Learning Flows in Digital Learning Ecosystems	2015
A Video Summarization Method Based on Spectral Clustering	2013
Category-Specific Video Summarization	2014
Empirical Evidence for Context-aware Interfaces to Pedestrian Navigation Systems	2014
Event-Oriented Semantic Data Generation for Medical Guidelines	2014
The design of smart educational environments	2016
An optimization methodology for machine learning strategies and regression problems in ballistic impact scenarios	2012
Supporting Citizen Inquiry: An Investigation of Moon Rock	2013
Preliminary numerical investigations of conformal predictors based on fuzzy logic classifiers	2015
Finding top cliques for keyword search from graphs in polynomial delay	2015
Sources of World Knowledge	2012
Strategic Sourcing of R&D: The Determinants of Success	2010
Graphical Modelling in Genetics and Systems Biology	2015
Comfort as a Multidimensional Preference Model for Energy Efficiency Control Issues	2012
Towards a Reference Mission Map for Performance Measurement in Humanitarian Supply Chains	2010
Lessons Learnt in Providing Product Designers with User-Participatory Interaction Design Tools	2009
Introduction	2016
An Interactional View of Context in Business Processes	2014
A Semantic Model for Enhancing Network Services Management and Auditing	2011
Chapter 6. Implementation Aspects	2010
Coastal Monitoring: A Methodological Proposal for New Generation Coastal Planning in Apulia	2015
The Impact of Parameterized Complexity to Interdisciplinary Problem Solving	2012
Evaluation of Collaborative Enterprises Networks: Case Study of Brazilian Virtual Enterprises	2010
Efficient supervised and semi-supervised approaches for affiliations disambiguation	2013
Effectiveness and efficiency of a domain-specific language for high-performance marine ecosystem simulation: a controlled experiment	2016
Comprehending a Service by Informative Models	2016
The Impact of Organizational Learning on Innovativeness in Spanish Companies	2009
Evaluation Measures for TCBR Systems	2008
Evaluation Constructs for Visual Video Summaries	2010
Abstraction and Generalization in Reinforcement Learning: A Summary and Framework	2010
Link Analysis and Web Search	2012
Key Performance Indicators in Data Warehouses	2016
Low reciprocity rates in acquaintance networks of young adults: fact or artifact?	2014
Bidirectional Transformations: A Cross-Discipline Perspective	2009
Morphological mismatches in machine translation	2008
Acceptance of Mobile Entertainment by Chinese Rural People	2009
Reaping the benefits of big data in telecom	2016
Ontological reasoning for improving the treatment of emotions in text	2010
A Mobile Course Coordinator System	2011
Automatic Morphological Query Expansion Using Analogy-Based Machine Learning	2007
Finding related sentence pairs in MEDLINE	2010
Chapter 8: Multimedia and Multimodal Information Retrieval	2010
Ontological Reasoning to Configure Emotional Voice Synthesis	2007
A parallel evolutionary algorithm for technical market indicators optimization	2013
Personalizing Web Search Results Based on Subspace Projection	2014
Educational Ergonomics and the Future of the Mind	2016
Coupled Evolution of Software Metamodels and Models	2014
Towards Value-Oriented Use of Social Media for Knowledge Management in SME	2013
Integrating User Experience into a Software Development Company – A Case Study	2009
Role of informatics in the ecologization of society	2012
On the dataset shift problem in software engineering prediction models	2012
"No shit" or "Oh, shit!": responses to observations on the use of UML in professional practice	2014
Studying Evolving Software Ecosystems based on Ecological Models	2014
Contract Clarity and Usability through Visualization	2013
SVI: A Simple Single-Nucleotide Human Variant Interpretation Tool for Clinical Use	2015
Website Credibility, Active Trust and Behavioural Intent	2008
Computer Forensics and Culture	2008
Multi-hierarchy Information Visualization Research Based on Three-Dimensional Display of Products System	2009
A Structural Equation Model of Knowledge Management Practices and Library Users' Satisfaction at Malaysian University Libraries	2014
Digital Geometry and Its Applications to Medical Imaging	2009
Coupling Surveys with GPS Tracking to Explore Tourists' Spatio-Temporal Behaviour	2016

A.2 Results of the tool for SLR case study.

P_ID	Title	Quality Criteria Score							
		Keyword	Abstract	Author	Pages	Type	Publisher	Cited-by	Final
162	Kaiwa: Towards a method for knowledge transfer in the transition phase of Offshore outsourced projects	0.035	0.5	1	1	1	1	0.01	4.55
39	Exploring storytelling as a knowledge transfer technique in offshore outsourcing	0.04	0.5	1	1	1	1	0	4.54
98	An integrated model of success in IT outsourcing relationships: Implications for the public sector	0.028	0.5	1	1	1	1	0	4.53
175	From research to practical application: Knowledge transfer planning and execution in outsourcing	0.025	0.5	1	1	1	1	0	4.53
16	IT outsourcing in the public sector: A conceptual model	0.01	0.5	1	1	1	1	0.01	4.52
1041	Gaining from vertical partnerships: Knowledge transfer relationship duration and supplier performance improvement in the U.S. and Japanese automotive industries	0.02	0.5	0	1	1	1	1	4.52
1073	The impact of knowledge sharing organizational capability and partnership quality on IS outsourcing success	0.013	0.5	0	1	1	1	0.98	4.49
740	Global outsourcing of back office services: Lessons trends and enduring challenges	0.026	1	0	1	1	1	0.21	4.24
159	Knowledge sharing in offshore software development a vendor perspective	0.037	0.5	0.67	1	1	1	0.01	4.22
43	Methodology fit in offshoring software development projects	0.014	0.5	0.67	1	1	1	0	4.18
147	Knowledge transfer challenges and mitigation strategies in global software development - A systematic literature review and industrial validation	0.053	1	0	1	1	1	0.06	4.11
978	IT outsourcing success: A psychological contract perspective	0.02	0.5	0	1	1	1	0.59	4.11
997	IT outsourcing strategies: Universalistic contingency and configurational explanations of success	0.032	0.5	0	1	1	1	0.57	4.1
569	A business process outsourcing framework based on business process management and knowledge management	0.039	1	0	1	1	1	0.05	4.09
938	A comparison of leadership roles in internal IT projects versus outsourcing projects	0.015	1	0	1	1	1	0.07	4.09
264	Security risks in service offshoring and outsourcing	0.044	1	0	1	1	1	0.04	4.08
572	Knowledge work productivity in distributed teams	0.015	1	0	1	1	1	0.06	4.08
1058	Knowledge integration in virtual teams: The potential role of KMS	0.026	0.5	0	1	1	1	0.55	4.08
135	Empirical studies on the use of social software in global software development-A systematic mapping study	0.021	1	0	1	1	1	0.05	4.07
823	The use of knowledge management in software acquisition	0.046	1	0	1	1	1	0.02	4.07
842	Empirical study of interactions between knowledge management activities	0.013	1	0	1	1	1	0.05	4.06
25	Knowledge4Scrum a novel knowledge management tool for agile distributed teams	0.049	1	0	1	1	1	0	4.05
442	Offshore supplier relations: knowledge integration among small businesses	0.03	1	0	1	1	1	0.02	4.05
626	Outsourcing as a mode of organizational learning	0.015	1	0	1	1	1	0.03	4.05
714	Benchmarking knowledge gaps through role simulations for assessing outsourcing viability	0.039	1	0	1	1	1	0.01	4.05
744	The distributed team members' explanations of knowledge they assume to be shared	0.02	1	0	1	1	1	0.02	4.04
901	Outsourcing of public services and implications for managerial knowledge and careers	0.02	1	0	1	1	1	0.01	4.03
6	Vendors' team performance in software outsourcing projects: From the perspective of transactive memory systems behavioral characteristics	0.021	1	0	1	1	1	0	4.02
85	Outsourcing of knowledge processes: A literature review	0.02	1	0	1	1	1	0	4.02
482	How knowledge and technology relate in creating value: An Italian case of technology outsourcing	0.018	1	0	1	1	1	0	4.02
843	Under what conditions do subsidiaries learn?	0.019	1	0	1	1	1	0	4.02
121	Establishing trust in Offshore software outsourcing relationships: An exploratory study using a systematic literature review	0.03	0.5	0.33	1	1	1	0.04	3.9
692	Exploring the effects of trust, task interdependence and virtualness on knowledge sharing in teams	0.025	0.5	0	1	1	1	0.37	3.9
30	Messy talk in virtual teams: Achieving knowledge synthesis through shared visualizations	0.045	0.5	0.33	1	1	1	0.01	3.89
64	Coordinating expertise across knowledge boundaries in offshore-outsourcing projects: The role of codification	0.01	0.5	0.33	1	1	1	0.05	3.89
202	Knowledge sharing management risks in outsourcing from various continents perspective: A systematic literature review	0.038	0.5	0.33	1	1	1	0	3.87
962	Perceived individual collaboration know-how development through information technology-enabled contextualization: Evidence from distributed teams	0.022	0.5	0	1	1	1	0.34	3.86

P_ID	Title	Quality Criteria Score							
32	Messy work in virtual worlds: exploring discovery and synthesis in virtual teams	0.013	0.5	0.33	1	1	1	0	3.84
717	Transformational technologies and the creation of new work practices: Making implicit knowledge explicit in task-based offshoring	0.055	0.5	0	1	1	1	0.26	3.82
691	Knowledge transfer in globally distributed teams: The role of transactive memory	0.041	0.5	0	1	1	1	0.27	3.81
749	An ontology-based approach to knowledge management in design processes	0.026	0.5	0	1	1	1	0.27	3.8
463	Global software development and collaboration: Barriers and solutions	0.048	0.5	0	1	1	1	0.23	3.78
1008	It and business process outsourcing: The knowledge potential	0.011	0.5	0	1	1	1	0.27	3.78
839	Does peripheral knowledge complement control? An empirical test in technology outsourcing alliances	0.015	0.5	0	1	1	1	0.22	3.74
966	A knowledge-based supplier intelligence retrieval system for outsource manufacturing	0.017	0.5	0	1	1	1	0.22	3.74
1003	Trust-building mechanisms utilized in outsourced IS development projects: A case study	0.028	0.5	0	1	1	1	0.21	3.74
33	Knowledge sharing in distributed agile projects: Techniques, strategies and challenges	0.04	0.5	0.67	1	1	0.5	0.01	3.72
578	Knowledge transfer processes in IT outsourcing relationships and their impact on shared knowledge and outsourcing performance	0.052	0.5	0	1	1	1	0.17	3.72
956	Utilizing knowledge context in virtual collaborative work	0.006	0.5	0	1	1	1	0.21	3.72
100	IT tools for knowledge storage and retrieval in globally distributed complex software and systems development of high-tech organizations	0.036	0.5	0.67	1	1	0.5	0	3.71
101	Knowledge sharing and application in complex software and systems development in globally distributed high-tech organizations using suitable IT tools	0.039	0.5	0.67	1	1	0.5	0	3.71
840	A comparison of transaction cost, agency and knowledge-based theory predictors of IT outsourcing decisions: A U.S.-Japan cross-cultural field study	0.012	0.5	0	1	1	1	0.2	3.71
993	Enabling knowledge creation in far-flung teams: Best practices for IT support and knowledge sharing	0.02	0.5	0	1	1	1	0.19	3.71
733	Transformational offshore outsourcing: Empirical evidence from alliances in China	0.038	0.5	0	1	1	1	0.16	3.7
854	Knowledge sharing and cooperation in outsourcing projects - A game theoretic analysis	0.009	0.5	0	1	1	1	0.19	3.7
899	Managing knowledge in global software development efforts: Issues and practices	0.033	0.5	0	1	1	1	0.17	3.7
1088	Computer-Mediated Inter-Organizational Knowledge-Sharing: Insights from a Virtual Team Innovating Using a Collaborative Toolsmall	0.008	0.5	0	1	1	1	0.19	3.7
604	The impact of outsourcing new technologies on integrative capabilities and performance	0.025	0.5	0	1	1	1	0.16	3.69
727	An integrative model of trust on IT outsourcing: Examining a bilateral perspective	0.022	0.5	0	1	1	1	0.17	3.69
959	Does being R&D intensive still discourage outsourcing?: Evidence from Dutch manufacturing	0.008	0.5	0	1	1	1	0.18	3.69
716	Managing the knowledge supply chain: An organizational learning model of information technology offshore outsourcing	0.02	0.5	0	1	1	1	0.16	3.68
1071	A design knowledge management system to support collaborative information product evolution	0.018	0.5	0	1	1	1	0.16	3.68
912	Collaborative activities in virtual settings: A knowledge management perspective of telemedicine	0.015	0.5	0	1	1	1	0.15	3.67
1040	A knowledge-based risk assessment framework for evaluating web-enabled application outsourcing projects	0.031	0.5	0	1	1	1	0.14	3.67
838	Knowledge management in 21st century manufacturing	0.014	0.5	0	1	1	1	0.15	3.66
951	IS outsourcing management competence dimensions: Instrument development and relationship exploration	0.01	0.5	0	1	1	1	0.15	3.66
144	Trust and knowledge sharing in diverse global virtual teams	0.028	0.5	0	1	1	1	0.12	3.65
243	Tools used in Global Software Engineering: A systematic mapping review	0.046	0.5	0	1	1	1	0.1	3.65
1070	Vertical technology transfer via international outsourcing	0.013	0.5	0	1	1	1	0.14	3.65
984	Technological outsourcing and product diversification: Do markets for technology affect firms' strategies?	0.007	0.5	0	1	1	1	0.13	3.64
991	Making organisations virtual: The hidden cost of distributed teams	0.015	0.5	0	1	1	1	0.11	3.63
418	Strategic orientations, knowledge acquisition and firm performance: The perspective of the vendor in cross-border outsourcing	0.021	0.5	0	1	1	1	0.1	3.62
493	Knowledge transfer processes for different experience levels of knowledge recipients at an offshore technical support center	0.055	0.5	0	1	1	1	0.06	3.62
693	Does interfirm modularity complement ignorance? A field study of software outsourcing alliances	0.018	0.5	0	1	1	1	0.1	3.62

P_ID	Title	Quality Criteria Score							
1000	Predicting information systems outsourcing success using a hierarchical design of case-based reasoning	0.031	0.5	0	1	1	1	0.09	3.62
1053	Knowledge flow management for distributed team software development	0.019	0.5	0	1	1	1	0.1	3.62
505	Organizational learning and capabilities for onshore and offshore business process outsourcing	0.015	0.5	0	1	1	1	0.09	3.61
217	Information technology outsourcing, knowledge transfer and firm productivity: An empirical analysis	0.024	0.5	0	1	1	1	0.08	3.6
355	Client-vendor knowledge transfer in IS offshore outsourcing: Insights from a survey of Indian software engineers	0.037	0.5	0	1	1	1	0.06	3.6
582	Offshore outsourcing in global design networks	0.028	0.5	0	1	1	1	0.07	3.6
908	From business process outsourcing (BPO) to knowledge process outsourcing (KPO): Some issues	0.021	0.5	0	1	1	1	0.08	3.6
954	An outsourcing decision model for sustaining long-term performance	0.025	0.5	0	1	1	1	0.07	3.6
136	When global virtual teams share knowledge: Media richness, cultural difference and language commonality	0.014	0.5	0	1	1	1	0.08	3.59
338	Knowledge sharing through virtual teams across borders and boundaries	0.035	0.5	0	1	1	1	0.05	3.59
354	Cognitive conflict and consensus generation in virtual teams during knowledge capture: Comparative effectiveness of techniques	0.049	0.5	0	1	1	1	0.04	3.59
364	Traceability and management of dispersed product knowledge during design and manufacturing	0.024	0.5	0	1	1	1	0.07	3.59
371	Effects of initial and ongoing trust in IT outsourcing: A bilateral perspective	0.019	0.5	0	1	1	1	0.07	3.59
376	The role of IT human capability in the knowledge transfer process in IT outsourcing context	0.023	0.5	0	1	1	1	0.07	3.59
579	The impact of virtual technologies on knowledge-based processes: An empirical study	0.021	0.5	0	1	1	1	0.07	3.59
739	Offshoring technology innovation: A case study of rare-earth technology	0.011	0.5	0	1	1	1	0.08	3.59
14	Knowledge transfer and utilization in IT outsourcing partnerships: A preliminary model of antecedents and outcomes	0.023	0.5	0	1	1	1	0.06	3.58
472	Determinants of success in IS offshoring projects: Results from an empirical study of German companies	0.016	0.5	0	1	1	1	0.06	3.58
760	Questionnaire-based risk assessment scheme for Japanese offshore software outsourcing	0.03	0.5	0	1	1	1	0.05	3.58
829	24-hour knowledge factory: Using Internet technology to leverage spatial and temporal separations	0.022	0.5	0	1	1	1	0.06	3.58
851	Design complexity, vertical disintegration and knowledge organization in the semiconductor industry	0.012	0.5	0	1	1	1	0.07	3.58
927	Task partitioning in new product development teams: A knowledge and learning perspective	0.011	0.5	0	1	1	1	0.07	3.58
24	The mediating effect of knowledge sharing on the relationship between trust and virtual team effectiveness	0.025	0.5	0	1	1	1	0.04	3.57
138	KM and Global Software Engineering (GSE)	0.058	0.5	0	1	1	1	0.01	3.57
237	Reference architecture, metamodel and modeling principles for architectural knowledge management in information technology services	0.037	0.5	0	1	1	1	0.03	3.57
320	Exploring the motives and determinants of innovation performance of Malaysian offshore international joint ventures	0.01	0.5	0	1	1	1	0.06	3.57
370	A dynamic model of offshore software development	0.028	0.5	0	1	1	1	0.04	3.57
451	Perceived job effectiveness in coopetition: A survey of virtual teams within business organizations	0.013	0.5	0	1	1	1	0.06	3.57
470	Exploring the effects of online social ties on knowledge sharing: A comparative analysis of collocated vs dispersed teams	0.028	0.5	0	1	1	1	0.04	3.57
712	Exploratory study on effective Control structure in global business process sourcing	0.04	0.5	0	1	1	1	0.03	3.57
1086	Knowledge management as a framework for understanding public sector outsourcing	0.03	0.5	0	1	1	1	0.04	3.57
1087	Knowledge acquisition in virtual teams	0.015	0.5	0	1	1	1	0.05	3.57
127	Acquiring and Sharing tacit knowledge in software development teams: An empirical study	0.024	0.5	0	1	1	1	0.04	3.56
246	Global outsourcing relationships and innovation: A conceptual framework and research propositions	0.025	0.5	0	1	1	1	0.03	3.56
250	Knowledge management: A Solution to requirements understanding in global software engineering	0.061	0.5	0	1	1	1	0	3.56
315	Understanding transition performance during offshore IT outsourcing	0.03	0.5	0	1	1	1	0.03	3.56
366	Knowledge transfer in offshore outsourcing: A case study of Japanese and vietnamese software companies	0.036	0.5	0	1	1	1	0.02	3.56
368	Modeling job effectiveness and its antecedents from a social capital perspective: A survey of virtual teams within business organizations	0.026	0.5	0	1	1	1	0.03	3.56
469	Virtual software team project management	0.02	0.5	0	1	1	1	0.04	3.56
497	How to get mature global virtual teams: A framework to improve team process management in distributed software teams	0.032	0.5	0	1	1	1	0.03	3.56

P_ID	Title	Quality Criteria Score							
517	Software architecture evaluation in global software development projects	0.036	0.5	0	1	1	1	0.02	3.56
589	A Learning model of information technology outsourcing: Normative implications	0.016	0.5	0	1	1	1	0.04	3.56
597	Mapping knowledge flows in virtual teams with SNA	0.017	0.5	0	1	1	1	0.04	3.56
628	Knowledge outsourcing: An alternative strategy for knowledge management	0.026	0.5	0	1	1	1	0.03	3.56
646	A lightweight approach for knowledge sharing in distributed software teams	0.015	0.5	0	1	1	1	0.04	3.56
658	Knowledge-sharing in crossfunctional virtual teams	0.021	0.5	0	1	1	1	0.04	3.56
808	Vague knowledge search in the design for outsourcing using fuzzy decision tree	0.017	0.5	0	1	1	1	0.04	3.56
913	Research propositions for knowledge management systems supporting IT outsourcing relationships	0.015	0.5	0	1	1	1	0.04	3.56
1034	Experiences from knowledge management implementations in companies of the software sector	0.016	0.5	0	1	1	1	0.04	3.56
10	Internalization of R&D outsourcing: An empirical study	0.016	0.5	0	1	1	1	0.03	3.55
111	Knowledge transfer and knowledge building at offshored technical support centers	0.024	0.5	0	1	1	1	0.03	3.55
205	The impact of team task and job engagement on the transfer of tacit knowledge in e-business virtual teams	0.019	0.5	0	1	1	1	0.03	3.55
239	The contribution of in-house and external design activities to product market performance	0.011	0.5	0	1	1	1	0.04	3.55
339	Factors impacting knowledge transfer success in information systems outsourcing	0.018	0.5	0	1	1	1	0.03	3.55
405	Cross-cultural knowledge management practices to support offshore outsourcing	0.035	0.5	0	1	1	1	0.01	3.55
518	Measuring knowledge creation in virtual teams through the social network analysis	0.011	0.5	0	1	1	1	0.04	3.55
630	Working in distributed teams: Challenges, best practices and guidelines	0.048	0.5	0	1	1	1	0	3.55
766	Evaluating collaboration platforms for offshore software development scenarios	0.026	0.5	0	1	1	1	0.02	3.55
858	Knowledge transfer in IT outsourcing relationships: Three international case studies	0.01	0.5	0	1	1	1	0.04	3.55
948	Knowledge management in global software teams	0.027	0.5	0	1	1	1	0.02	3.55
9	Inter-firm learning and knowledge-sharing in multinational networks: An outsourced organization's perspective	0.012	0.5	0	1	1	1	0.03	3.54
13	Managing global software projects	0.035	0.5	0	1	1	1	0	3.54
18	Navigating the mutual knowledge problem: A comparative case study of distributed work	0.026	0.5	0	1	1	1	0.01	3.54
76	Relatedness effect on it outsourcing strategies and knowledge management on organizational performance impact on Government Papua Province	0.042	0.5	0	1	1	1	0	3.54
157	Key factors in managing IT outsourcing relationships	0.027	0.5	0	1	1	1	0.01	3.54
164	The cross-cultural knowledge sharing challenge: An investigation of the co-location strategy in software development offshoring	0.035	0.5	0	1	1	1	0	3.54
245	Knowledge transfer to vendors in offshore information systems outsourcing: Antecedents and effects on performance	0.025	0.5	0	1	1	1	0.01	3.54
258	Building socioemotional environments in metaverses for virtual teams in healthcare: A conceptual exploration	0.025	0.5	0	1	1	1	0.01	3.54
273	Transfer of method knowledge and modelling in distributed teams - Lessons learned	0.041	0.5	0	1	1	1	0	3.54
279	The key role of interfaces in IT outsourcing relationships	0.033	0.5	0	1	1	1	0.01	3.54
297	Occupational stress, knowledge sharing and GSD communication barriers as predictors of software engineer's creativity	0.04	0.5	0	1	1	1	0	3.54
318	Determinants of software quality in offshore development - An empirical study of an Indian vendor	0.012	0.5	0	1	1	1	0.03	3.54
327	Supporting communication and cooperation in global software development with agile service networks	0.025	0.5	0	1	1	1	0.01	3.54
491	Conceptualising aerospace outsourcing: Airbus UK and the lean supply approach	0.017	0.5	0	1	1	1	0.02	3.54
593	Knowledge transfer success factors in IT outsourcing environment	0.027	0.5	0	1	1	1	0.01	3.54
707	Risk bias externalization for offshore software outsourcing by conjoint analysis	0.031	0.5	0	1	1	1	0.01	3.54
713	Analyzing IT maintenance outsourcing decision from a knowledge management perspective	0.012	0.5	0	1	1	1	0.03	3.54
732	The knowledge advantage of virtual teams Processes supporting knowledge synergy	0.017	0.5	0	1	1	1	0.02	3.54
743	Transaction costs and organisational learning in strategic outsourcing relationships	0.014	0.5	0	1	1	1	0.03	3.54
751	"""kt"" carepacks -collaboration patterns for knowledge transfer: Learning from is/itoutsourcing case at a Swiss financial institution"	0.035	0.5	0	1	1	1	0	3.54
821	Sharing knowledge in global virtual teams: How do Chinese team members perceive the impact of national cultural differences on knowledge sharing?	0.025	0.5	0	1	1	1	0.01	3.54

P_ID	Title	Quality Criteria Score							
861	Knowledge acquisition and risk analysis in material recovery facilities by a virtual team	0.043	0.5	0	1	1	1	0	3.54
873	Knowledge management in offshore software development	0.027	0.5	0	1	1	1	0.01	3.54
910	From compliance to business success: Improving outsourcing service controls by adopting external regulatory requirements	0.031	0.5	0	1	1	1	0.01	3.54
915	Using Social Networking Analysis to Facilitate Knowledge Sharing in the British Council	0.011	0.5	0	1	1	1	0.03	3.54
967	Siemens: Expanding the knowledge management system sharenet to research & development	0.017	0.5	0	1	1	1	0.02	3.54
1019	Knowledge integration in new product development: The FIAT autocase	0.028	0.5	0	1	1	1	0.01	3.54
1059	Outsourcing facilities management in the process industry: A comparison of Swedish and UK patterns	0.011	0.5	0	1	1	1	0.03	3.54
3	Strategic outsourcing under technology spillovers	0.02	0.5	0	1	1	1	0.01	3.53
37	Knowledge processes in virtual teams: Consolidating the evidence	0.023	0.5	0	1	1	1	0.01	3.53
68	Choice of supply chain governance: Self-managing or outsourcing?	0.022	0.5	0	1	1	1	0.01	3.53
78	Outsourcing versus technology transfer: Hotelling meets Stackelberg	0.015	0.5	0	1	1	1	0.01	3.53
140	Investigating the roles of interpersonal and interorganizational trust in IT outsourcing success	0.016	0.5	0	1	1	1	0.01	3.53
154	Knowledge acquisition by outsourced service providers from aging workforce of oil and gas industry: A study	0.026	0.5	0	1	1	1	0	3.53
224	Towards a framework for transferring technology knowledge between facilities	0.015	0.5	0	1	1	1	0.01	3.53
260	Exploration of social capital and knowledge sharing: An empirical study on student virtual teams	0.023	0.5	0	1	1	1	0.01	3.53
268	Enhancing international knowledge transfer through information technology: The intervention of communication culture	0.018	0.5	0	1	1	1	0.01	3.53
274	Learning software-maintenance tasks in the transition phase of offshore outsourcing projects: Two learning-theoretical perspectives	0.03	0.5	0	1	1	1	0	3.53
276	Research article the role of leadership and contextualization on citizenship behaviors in distributed teams: A relational capital perspective	0.017	0.5	0	1	1	1	0.01	3.53
292	Supporting Cross-Border knowledge transfer through Virtual Teams communities and ICT Tools	0.024	0.5	0	1	1	1	0.01	3.53
314	An ontology-based method for measurement of transferability and complexity of knowledge in multi-site software development environment	0.024	0.5	0	1	1	1	0.01	3.53
337	Managing emergent knowledge through deferred action design principles: The case of ecommerce virtual teams	0.028	0.5	0	1	1	1	0	3.53
443	Technology transfer of dynamic IT outsourcing requires security measures in SLAs	0.016	0.5	0	1	1	1	0.01	3.53
487	Action readiness and mindset for IT offshoring	0.014	0.5	0	1	1	1	0.02	3.53
635	Measuring the risks of outsourcing: Experiences from industry	0.032	0.5	0	1	1	1	0	3.53
709	Reflection on knowledge sharing in F/OSS projects	0.016	0.5	0	1	1	1	0.01	3.53
710	How information systems providers develop and manage expertise and leverage their client relationships for competitive advantage	0.026	0.5	0	1	1	1	0	3.53
764	Offshore software development: Transferring research findings into the classroom	0.022	0.5	0	1	1	1	0.01	3.53
922	Enabling knowledge creation in far-flung teams: Best practices for IT support and knowledge sharing	0.026	0.5	0	1	1	1	0	3.53
937	Challenges in developing a knowledge management strategy for the air force material command	0.019	0.5	0	1	1	1	0.01	3.53
975	Effective knowledge transfer in virtual teams: Linking contents and mechanisms	0.029	0.5	0	1	1	1	0	3.53
1018	Knowledge transfer through the supply system: Does modularity make it easier?	0.008	0.5	0	1	1	1	0.02	3.53
11	Controlled experiments as means to teach soft skills in software engineering	0.022	0.5	0	1	1	1	0	3.52
22	Co-creating the work structures and processes of the future: The interplay of virtual teams and knowledge management	0.017	0.5	0	1	1	1	0	3.52
83	Strategic outsourcing with technology transfer under cournot competition	0.008	0.5	0	1	1	1	0.01	3.52
113	Developing a web-based knowledge product outsourcing system at a university	0.012	0.5	0	1	1	1	0.01	3.52
118	Software architecture in distributed software development: A review	0.017	0.5	0	1	1	1	0	3.52
150	Understanding the role of representation in interorganizational knowledge integration: A case study of an IT outsourcing project	0.022	0.5	0	1	1	1	0	3.52
216	Cultural fit, knowledge transfer and innovation performance: a study of Malaysian offshore international joint ventures	0.014	0.5	0	1	1	1	0.01	3.52

P_ID	Title	Quality Criteria Score							
267	Managing knowledge in internationalizing universities through foreign assignments	0.013	0.5	0	1	1	1	0.01	3.52
322	Trusty: A tool to improve communication and collaboration in DSD	0.024	0.5	0	1	1	1	0	3.52
328	Measurement of the development of a learning IT organization supported by a model of knowledge acquisition and processing	0.022	0.5	0	1	1	1	0	3.52
347	Knowledge Integration Challenges when Outsourcing Manufacturing	0.022	0.5	0	1	1	1	0	3.52
457	The 24-hour knowledge factory: Work and organizational redesign and associated challenges	0.02	0.5	0	1	1	1	0	3.52
502	Parallel transitions in IT outsourcing: Making it happen	0.018	0.5	0	1	1	1	0	3.52
514	Toward a general model of knowledge creation in virtual teams	0.023	0.5	0	1	1	1	0	3.52
627	Implementing global virtual teams to enhance cross-border transfer of knowledge in multinational enterprises: A resource-based view	0.011	0.5	0	1	1	1	0.01	3.52
729	Roadmap for knowledge sharing and transfer: Sustaining outsourcing relationships	0.011	0.5	0	1	1	1	0.01	3.52
769	Effects of social identity processes on coordination and knowledge sharing in geographically distributed software teams	0.015	0.5	0	1	1	1	0	3.52
779	Small firms and offshore software outsourcing: High transaction costs and their mitigation	0.016	0.5	0	1	1	1	0	3.52
828	Impact of Internet-based distributed monitoring systems on offshore sourcing of services	0.019	0.5	0	1	1	1	0	3.52
890	Business process outsourcing to emerging markets: A knowledge management approach to models and strategies	0.015	0.5	0	1	1	1	0	3.52
924	Leading Together, Working Together: The Role of Team Shared Leadership in Building Collaborative Capital in Virtual Teams	0.008	0.5	0	1	1	1	0.01	3.52
929	Forces affecting offshore software development	0.011	0.5	0	1	1	1	0.01	3.52
934	Multi-agent support for distributed engineering design	0.006	0.5	0	1	1	1	0.01	3.52
1027	An inside look at outsourcing	0.015	0.5	0	1	1	1	0	3.52
4	Roles of culture for knowledge sharing in asian virtual teams: A case study	0.011	0.5	0	1	1	1	0	3.51
29	Transactive Memory System, Communication Quality, and Knowledge Sharing in Distributed Teams: An Empirical Examination in Open Source Software Project Teams	0.012	0.5	0	1	1	1	0	3.51
84	Knowledge management: A cross sectorial comparison of wind generation and naval engineering	0.012	0.5	0	1	1	1	0	3.51
99	A self-learning approach to improving service quality in outsourcing of engineering design using operational data	0.013	0.5	0	1	1	1	0	3.51
128	Enhancing quality of it services delivery using enterprise crowdsourcing	0.013	0.5	0	1	1	1	0	3.51
133	Vertical FDI versus outsourcing: The role of technology transfer costs	0.013	0.5	0	1	1	1	0	3.51
190	Developing professional competencies using a Living Lab approach: An exploratory study in the field of management education	0.011	0.5	0	1	1	1	0	3.51
348	Trade-offs in Make-Buy Decisions: Exploring Operating Realities of Knowledge Integration and Innovation	0.009	0.5	0	1	1	1	0	3.51
523	Managing the information technology: Knowledge transfer in virtual teams	0.006	0.5	0	1	1	1	0	3.51
571	Knowledge value creation characteristics of virtual teams: A case study in the construction sector	0.008	0.5	0	1	1	1	0	3.51
660	The possibility of water-cooler chat? Developing communities of practice for knowledge sharing within global virtual teams	0.011	0.5	0	1	1	1	0	3.51
750	Knowledge acquisition within a network of industrial laboratories performing quality specifications testing	0.013	0.5	0	1	1	1	0	3.51
754	Sourcing decisions in the finance sector: current state and focus on economies of scale and knowledge management	0.009	0.5	0	1	1	1	0	3.51
822	It-outsourcing and it-offshoring: Trends and impacts on SE/KE curricula	0.011	0.5	0	1	1	1	0	3.51
1063	Developing collective knowledge in geographically distributed teams	0.008	0.5	0	1	1	1	0	3.51
117	A survey of tools for mapping and synchronization of knowledge from legacy systems	0.023	0.5	0.33	1	1	0.5	0	3.35
214	A conceptual framework of knowledge integration in multisourcing arrangements	0.019	0.5	0.33	1	1	0.5	0	3.35
872	A research agenda for distributed software development	0.017	0.5	0	1	1	0.5	0.25	3.27
1065	The impact of knowledge sharing organizational capability and partnership quality on IS outsourcing success	0.018	0.5	0	1	1	0.5	0.21	3.23
1046	An insight into the interplay between culture, conflict and distance in globally distributed requirements negotiations	0.019	0.5	0	1	1	0.5	0.13	3.15

P_ID	Title	Quality Criteria Score							
565	Knowledge management in distributed software development teams - does culture matter?	0.024	0.5	0	1	1	0.5	0.07	3.09
149	Factors affecting the outcome of Global Software Development projects: An empirical study	0.039	0.5	0	1	1	0.5	0.02	3.06
352	An analysis of collaborative patterns in large-scale ontology development projects	0.02	0.5	0	1	1	0.5	0.04	3.06
844	Cultural influences and globally distributed information systems development: Experiences from Chinese IT professionals	0.012	0.5	0	1	1	0.5	0.05	3.06
340	Experiences with semantic wikis for architectural knowledge management	0.019	0.5	0	1	1	0.5	0.03	3.05
341	Architecturally significant requirements, reference architecture and metamodel for knowledge management in information technology services	0.04	0.5	0	1	1	0.5	0.01	3.05
385	Using wikis to support architectural knowledge management in global software development	0.032	0.5	0	1	1	0.5	0.02	3.05
540	Investigating knowledge transfer mechanisms for oil rigs	0.049	0.5	0	1	1	0.5	0	3.05
793	Observing software testing practice from the viewpoint of organizations and knowledge management	0.03	0.5	0	1	1	0.5	0.02	3.05
158	An ontology for task allocation to teams in distributed software development	0.022	0.5	0	1	1	0.5	0.02	3.04
196	Knowledge management for services transitions	0.036	0.5	0	1	1	0.5	0	3.04
212	Multi-level knowledge transfer in software development outsourcing projects: The agency theory view	0.018	0.5	0	1	1	0.5	0.02	3.04
321	Adapting business and technical processes for collaborative software development	0.034	0.5	0	1	1	0.5	0.01	3.04
480	Exploring the role of Transactive Memory System (TMS) for knowledge transfer processes in Malaysia E-government IT outsourcing	0.036	0.5	0	1	1	0.5	0	3.04
520	Critical issues of offshore software development project failures	0.03	0.5	0	1	1	0.5	0.01	3.04
535	Factors affecting the willingness of multinational corporation's to share knowledge with universities: A case study of the automotive industry in Thailand	0.038	0.5	0	1	1	0.5	0	3.04
542	Leveraging semantic data wikis for distributed requirements elicitation	0.018	0.5	0	1	1	0.5	0.02	3.04
552	Multi-vendor outsourcing: Relational structures and organizational learning from a social relation perspective	0.016	0.5	0	1	1	0.5	0.02	3.04
674	A collaborative platform for application knowledge management in software maintenance projects	0.008	0.5	0	1	1	0.5	0.03	3.04
763	Successful collaborative software projects for medical devices in an FDA regulated environment: Myth or reality?	0.031	0.5	0	1	1	0.5	0.01	3.04
771	The role of IT human capability in knowledge transfer process under IT outsourcing situations	0.025	0.5	0	1	1	0.5	0.01	3.04
865	Collaborative online training methods in global production flow	0.043	0.5	0	1	1	0.5	0	3.04
880	From telco to techno: Outsourcing transitioning in information systems	0.035	0.5	0	1	1	0.5	0	3.04
1044	Knowledge transfer in virtual information systems development teams: An empirical examination of key enablers	0.011	0.5	0	1	1	0.5	0.03	3.04
47	Why closely coupled work matters in global software development	0.032	0.5	0	1	1	0.5	0	3.03
59	Knowledge sharing, social capital and firm performance in technological clusters of Taiwan Science Parks: An innovation strategy perspective	0.03	0.5	0	1	1	0.5	0	3.03
109	Research on the multiple combination weight based on rough set and clustering analysis-the knowledge transfer risk in it outsourcing taken as an example	0.015	0.5	0	1	1	0.5	0.01	3.03
123	An empirical study of the relationship between team social capital and knowledge transfer: Mediating role of transactive memory system	0.025	0.5	0	1	1	0.5	0	3.03
124	Tacit knowledge acquisition in virtual teams	0.028	0.5	0	1	1	0.5	0	3.03
223	Fostering cross-site coordination through awareness: An investigation of state-of-the-practice through a focus group study	0.032	0.5	0	1	1	0.5	0	3.03
227	Tacit knowledge transfer through co-activation: A case study of design and support by an electronics manufacturing service firm	0.028	0.5	0	1	1	0.5	0	3.03
275	Enterprise modelling in distributed teams - Lessons learned from information demand modelling	0.031	0.5	0	1	1	0.5	0	3.03
295	Impact of knowledge management on offshore software development: An exploratory study	0.019	0.5	0	1	1	0.5	0.01	3.03
305	The influence of outsourcing models on vendor knowledge integration	0.025	0.5	0	1	1	0.5	0	3.03
306	Vendorvendor knowledge transfer in global ISD outsourcing projects: Insights from a German case study	0.019	0.5	0	1	1	0.5	0.01	3.03
519	Towards needdriven knowledge sharing in distributed teams	0.022	0.5	0	1	1	0.5	0.01	3.03
537	Knowledge process outsourcing	0.026	0.5	0	1	1	0.5	0	3.03
539	The effect of knowledge sharing on IS outsourcing success	0.019	0.5	0	1	1	0.5	0.01	3.03

P_ID	Title	Quality Criteria Score							
547	Antecedents of success in is offshoring projects - Proposal for an empirical research study	0.019	0.5	0	1	1	0.5	0.01	3.03
548	"QSHIP & advanced use of hydrodynamics in early design"	0.019	0.5	0	1	1	0.5	0.01	3.03
585	An analysis of knowledge outsourcing at Eduware	0.018	0.5	0	1	1	0.5	0.01	3.03
615	Determining a firm's optimal outsourcing rate: A learning model perspective	0.025	0.5	0	1	1	0.5	0	3.03
655	A study on the success factors of building regional contents industry clusters - Based on the Daegu Contents Industry	0.02	0.5	0	1	1	0.5	0.01	3.03
657	Virtual team common knowledge: Construct specification and effect on knowledge integration effectiveness	0.019	0.5	0	1	1	0.5	0.01	3.03
673	Process documentation, operational alignment and flexibility in it outsourcing relationships: A knowledge-based perspective	0.014	0.5	0	1	1	0.5	0.02	3.03
792	Requirements determination for knowledge management systems in information technology outsourcing relationships	0.025	0.5	0	1	1	0.5	0	3.03
794	Risks of outsourcing complex systems - A growing phenomenon	0.033	0.5	0	1	1	0.5	0	3.03
802	Contracting incentives and inter-organizational routines in Information Technology outsourcing	0.027	0.5	0	1	1	0.5	0	3.03
806	Implementing centralised IT service management: Drawing lessons from the public sector	0.012	0.5	0	1	1	0.5	0.02	3.03
815	Managing risks and maintaining a competitive edge in today's outsourcing environment	0.025	0.5	0	1	1	0.5	0	3.03
862	Social software as an enabler for innovations in e-collaboration	0.032	0.5	0	1	1	0.5	0	3.03
877	Knowledge transfer in offshore insourcing	0.014	0.5	0	1	1	0.5	0.02	3.03
935	An integrative model of trust on IT outsourcing : From the service receiver's perspective	0.022	0.5	0	1	1	0.5	0.01	3.03
1045	Supporting the knowledge life cycle with a knowledge network management system	0.02	0.5	0	1	1	0.5	0.01	3.03
63	Spanning knowledge boundaries in offshore insourcing through organizational learning: A case study of bankco	0.015	0.5	0	1	1	0.5	0	3.02
103	Investigating the impact of boundaries on knowledge sharing in global virtual teams	0.021	0.5	0	1	1	0.5	0	3.02
179	Knowledge and business intelligence technologies in cross-enterprise environments for Italian advanced mechanical industry - Project presentation	0.014	0.5	0	1	1	0.5	0.01	3.02
193	Psychological contract and knowledge management mediated by cultural dynamics	0.021	0.5	0	1	1	0.5	0	3.02
272	Transactive memory systems, knowledge integration and team performance in geographically dispersed teams	0.018	0.5	0	1	1	0.5	0	3.02
281	Understanding outsourcing commitment- An integrated model combining the resoruce-based view and knowledge management	0.018	0.5	0	1	1	0.5	0	3.02
330	Understanding organizational capabilities for effective offshoring	0.018	0.5	0	1	1	0.5	0	3.02
379	A study of successful outsourcing factors	0.023	0.5	0	1	1	0.5	0	3.02
393	Similarity and familiarity in distributed teams: A perspective of identification on knowledge sharing	0.022	0.5	0	1	1	0.5	0	3.02
406	Study on knowledge transfer in Knowledge Process Outsourcing (KPO)	0.016	0.5	0	1	1	0.5	0	3.02
409	Factors impacting knowledge transfer success in information systems outsourcing	0.011	0.5	0	1	1	0.5	0.01	3.02
416	Framework to address organizational gaps and build knowledge management capabilities in offshore KPO	0.022	0.5	0	1	1	0.5	0	3.02
424	Analysis about risk controlling in the application services based on knowledge management theory	0.017	0.5	0	1	1	0.5	0	3.02
428	Competing in the business process outsourcing industry: A call center case study	0.016	0.5	0	1	1	0.5	0	3.02
546	Institutional antecedents and consequences of modularity in business process outsourcing	0.019	0.5	0	1	1	0.5	0	3.02
551	Antecedents and consequences of modularization in BPO-based on transaction cost theory and knowledge-based theory	0.024	0.5	0	1	1	0.5	0	3.02
613	Encouraging knowledge sharing in global virtual teams: The interaction effect of individual difference and perceived sharing benefits	0.023	0.5	0	1	1	0.5	0	3.02
642	Management conditions influencing the successful outcome of IS outsourcing	0.016	0.5	0	1	1	0.5	0	3.02
671	The adoption of and satisfaction with Web2.0 based collaboration and knowledge management technologies in global software development - insights from an empirical study	0.017	0.5	0	1	1	0.5	0	3.02
758	Searching unanswered questions a review of knowledge management processes in virtual teams	0.022	0.5	0	1	1	0.5	0	3.02
759	Antecedents of knowledge sharing in globally distributed software development teams	0.02	0.5	0	1	1	0.5	0	3.02
799	Knowledge process outsourcing: Identifying potential research agenda based on industry trends	0.024	0.5	0	1	1	0.5	0	3.02

P_ID	Title	Quality Criteria Score							
805	Product development process managing in supply chain	0.022	0.5	0	1	1	0.5	0	3.02
864	Enabling offshore software testing: A case study	0.018	0.5	0	1	1	0.5	0	3.02
889	Critical success factors as a strategy for risk mitigation in IT outsourcing projects	0.02	0.5	0	1	1	0.5	0	3.02
926	Online knowledge sharing and media selection in a community organisation: An application of the theory of media synchronicity	0.02	0.5	0	1	1	0.5	0	3.02
940	Automated cleansing for spend analytics	0.01	0.5	0	1	1	0.5	0.01	3.02
36	Toward an integrative model of influence factors for success of global software development projects	0.012	0.5	0	1	1	0.5	0	3.01
161	Creating value through virtual teams: A current literature review	0.01	0.5	0	1	1	0.5	0	3.01
308	Offshoring attitudes, relational behaviours and departmental culture	0.007	0.5	0	1	1	0.5	0	3.01
392	The role of inspirational leadership and technology support for contextualization on psychological contract in distributed teams	0.012	0.5	0	1	1	0.5	0	3.01
415	Aiming at a moving target: IT alignment in toy companies	0.01	0.5	0	1	1	0.5	0	3.01
426	Integrated systems for product design: The move toward outsourcing	0.006	0.5	0	1	1	0.5	0	3.01
427	Vendor transition and the impact on in-flight projects	0.012	0.5	0	1	1	0.5	0	3.01
431	Deriving a research agenda for the management of multisourcing relationships based on a literature review	0.006	0.5	0	1	1	0.5	0	3.01
538	The influence of offshore development on strategies for developing China's software industry	0.012	0.5	0	1	1	0.5	0	3.01
553	Success factors for information systems outsourcing: A meta-analysis	0.011	0.5	0	1	1	0.5	0	3.01
631	From virtual teams to online communities: Fostering group based collaboration for innovation and knowledge management	0.008	0.5	0	1	1	0.5	0	3.01
652	Promotion of offshore software outsourcing enterprise value chain based on knowledge transfer effects	0.012	0.5	0	1	1	0.5	0	3.01
656	Study on Knowledge Management supporting virtual team operation	0.005	0.5	0	1	1	0.5	0	3.01
677	Assessing the impacts of is offshoring: Preliminary conclusions questioning the validity of cultural consideration	0.008	0.5	0	1	1	0.5	0	3.01
812	A comparative analysis of offshored and onshored software development projects	0.009	0.5	0	1	1	0.5	0	3.01
866	Outsourcing of ICT: An empirical study in Swiss SMEs	0.008	0.5	0	1	1	0.5	0	3.01
882	Inter-organizational knowledge development in IT outsourcing	0.014	0.5	0	1	1	0.5	0	3.01
883	AKSIO - Active knowledge management in the petroleum industry	0.01	0.5	0	1	1	0.5	0	3.01
928	Effect of cultural norms on media choice in Global Virtual Teams	0.011	0.5	0	1	1	0.5	0	3.01
213	Investigating the role of organizational structure in developing shared understanding of requirements within GSD	0.023	0.5	0.67	0	1	0.5	0	2.69
847	Outsourcing a core competency	0.017	0.5	0	0	1	1	0.03	2.55
734	Prototype of knowledge management system in Chinese offshore software development company	0.026	0.5	0	0	1	1	0.01	2.54
856	View from a business school: An interview with Professor Rita Marcella	0.006	0.5	0	0	1	1	0	2.51
197	Knowledge and contextual information management in global software development: Challenges and perspectives	0.041	0.5	0.33	0	1	0.5	0.01	2.38
28	Dedicated support for experience sharing in distributed software projects	0.033	0.5	0.33	0	1	0.5	0	2.36
444	Knowledge transfer in IT offshore outsourcing projects: An analysis of the current state and best practices	0.045	0.5	0	0	1	0.5	0.01	2.06
448	Knowledge transfer in global software development - Leveraging ontologies tools and assessments	0.045	0.5	0	0	1	0.5	0.01	2.06
668	Towards architectural knowledge management practices for global software development	0.022	0.5	0	0	1	0.5	0.04	2.06
795	ADkwik: Web 2.0 collaboration system for architectural decision engineering	0.027	0.5	0	0	1	0.5	0.03	2.06
165	Towards a global software development community web: Identifying patterns and scenarios	0.042	0.5	0	0	1	0.5	0.01	2.05
446	Architectural knowledge management in global software development: A review	0.022	0.5	0	0	1	0.5	0.03	2.05
449	Information security risk management: An empirical study on the importance and practices in ICT outsourcing	0.036	0.5	0	0	1	0.5	0.01	2.05
408	Process support for requirements engineering activities in global software development: A literature based evaluation	0.029	0.5	0	0	1	0.5	0.01	2.04
481	Information security risk factors: Critical threats and vulnerabilities in ICT outsourcing	0.032	0.5	0	0	1	0.5	0.01	2.04
664	Knowledge management system for screed design of asphalt paver	0.042	0.5	0	0	1	0.5	0	2.04

P_ID	Title	Quality Criteria Score							
1011	Virtual teams in knowledge-intensive and computer-mediated work	0.039	0.5	0	0	1	0.5	0	2.04
389	Research on model of knowledge transfer in outsourced software projects	0.021	0.5	0	0	1	0.5	0.01	2.03
515	Study on knowledge map construction for virtual team based on ontology	0.025	0.5	0	0	1	0.5	0	2.03
516	On educating globally distributed software development - A case study	0.027	0.5	0	0	1	0.5	0	2.03
530	Storytelling as a tool for knowledge transfer in the it industry	0.016	0.5	0	0	1	0.5	0.01	2.03
868	Knowledge transfer in system development offshore outsourcing projects	0.026	0.5	0	0	1	0.5	0	2.03
1057	Active artefact management for distributed software engineering	0.008	0.5	0	0	1	0.5	0.02	2.03
35	Exploring collaboration requirements of heterogeneous teams on the example of a critical sales process	0.016	0.5	0	0	1	0.5	0	2.02
75	"ITIL maturity model of IT outsourcing: Evidence from a ""leading user"""	0.015	0.5	0	0	1	0.5	0	2.02
204	Knowledge transfer in practice: A socio-technical system for the transition in outsourcing	0.019	0.5	0	0	1	0.5	0	2.02
219	Exploration of the evolution process of IT outsourcing from 1992 to 2011: A main path analysis	0.02	0.5	0	0	1	0.5	0	2.02
226	Integrated framework for virtual team management	0.019	0.5	0	0	1	0.5	0	2.02
299	Knowledge sharing management in offshore software development outsourcing relationships from vendors' perspective: A systematic literature review protocol	0.024	0.5	0	0	1	0.5	0	2.02
317	Knowledge management for product development: A review	0.024	0.5	0	0	1	0.5	0	2.02
334	CollectiveThought - Supporting virtual teams with semantic networks	0.021	0.5	0	0	1	0.5	0	2.02
460	The individual performance measurement framework in virtual team learning	0.019	0.5	0	0	1	0.5	0	2.02
536	Vendor's perspective on social capital and absorptive capacity for offshore contract performance	0.019	0.5	0	0	1	0.5	0	2.02
541	Research of risk evaluation of knowledge transfer in IT outsourcing based on RS-SVM	0.023	0.5	0	0	1	0.5	0	2.02
612	A model for enhancing knowledge, creation, application and succession while facilitating leadership change within virtual work environments	0.018	0.5	0	0	1	0.5	0	2.02
661	A new methodology for designing collaborative information system in a virtual team: A case study	0.015	0.5	0	0	1	0.5	0	2.02
680	Technology spillovers from international outsourcing: An empirical verification in China	0.017	0.5	0	0	1	0.5	0	2.02
1048	Software engineering education in developing countries - Responding to the outsourcing challenges	0.015	0.5	0	0	1	0.5	0	2.02
77	Contracting on project time in Knowledge Process outsourcing	0.013	0.5	0	0	1	0.5	0	2.01
283	A phenomenological study of the impact of knowledge intensity and environmental velocity on in-source or hosted contact centres	0.007	0.5	0	0	1	0.5	0	2.01
417	Knowledge characteristics and organizational learning strategies of service outsourcing enterprises	0.008	0.5	0	0	1	0.5	0	2.01
440	Managing effective knowledge acquisition in international outsourcing alliances	0.014	0.5	0	0	1	0.5	0	2.01
720	The impact of business process outsourcing on firm performance	0.014	0.5	0	0	1	0.5	0	2.01
1021	IS outsourcing management competence: Theoretical development and empirical examination	0.008	0.5	0	0	1	0.5	0	2.01

Bibliography

Ramon Abilio, Flávio Moraisand, Gustavo Vale, Claudiane Oliveira, Denilson Pereira, and Heitor Costa. Applying Information Retrieval Techniques to Detect Duplicates and to Rank References in the Preliminary Phases of Systematic Literature Reviews. *CLEI Electron. J.*, 18(2):1–24, 2015. (cited on Page 27, 29, and 31)

Sophia Ananiadou and Goran Nenadic. *Automatic terminology management in biomedicine*. Artech House, Boston/ London, 2006. (cited on Page 15)

Sophia Ananiadou, Brian Rea, Naoaki Okazaki, Rob Procter, and James Thomas. Supporting Systematic Reviews Using Text Mining. *Social Science Computer Review*, 27:470–300, 2009. (cited on Page 15 and 16)

Muhammad A. Babar and He Zhang. Systematic Literature Reviews in Software Engineering: Preliminary Results from Interviews with Researchers. In *Third International Symposium on Empirical Software Engineering and Measurement)*. IEEE, 2009. (cited on Page 1)

Sebastian K. Boell and Dubravka Cecez-Kecmanovic. On being 'systematic' in literature reviews in IS. *Journal of Information Technology (JIT)*, 30:161–173, 2015. (cited on Page 5 and 21)

Andrew Booth, Diana Papaioannou, and Anthea Sutton. *Systematic approaches to a successful literature review*. Sage publications, 2012. (cited on Page 5 and 6)

Pearl Brereton, Barbara A. Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. Lessons from applying the systematic literature review process within the software engineering domain. *The Journal of Systems and Software*, 80:571–583, 2007. (cited on Page 2)

Judy F. Burnham. Scopus database: a review. *Biomedical Digital Libraries*, 3, 2006. (cited on Page 49)

Fabio Q. B. da Silva, André L. M. Santos, Sérgio Soares, A. César C. França, Cleviton V. F. Monteiro, and Felipe F. Maciel. Six years of systematic literature reviews in software engineering: An updated tertiary study. *Information and Software Technology*, 53:899–913, 2011. (cited on Page 1 and 2)

Maya Daneva, Daniela Damianb, Alessandro Marchettoc, and Oscar Pastor. Empirical research methodologies and studies in Requirements Engineering: How far did we come? *The Journal of Systems and Software*, 94:1–9, 2014. (cited on Page 49)

Tore Dybå and Torgeir Dingsøyr. Strength of evidence in systematic reviews in software engineering. In *Empirical software engineering and measurement (ESEM)*, pages 178–187. ACM, 2008a. (cited on Page 2)

Tore Dybå and Torgeir Dingsøyr. Empirical studies of agile software development: A systematic review. *Information and Software Technology*, 50(9-10):833–859, 2008b. (cited on Page 25)

Tore Dybår, Torgeir Dingsøyr, and Geir K. Hanssen. Applying systematic reviews to diverse study types: An experience report. volume 7, pages 225–234. Empirical Software Engineering and Measurement, 2007. (cited on Page 6, 8, and 37)

Sandra Fabbri, Cleiton Silva, Elis Hernandes, Fábio Octaviano, André Di Thommazo, and Anderson Belgamo. Improvements in the StArt Tool to Better Support the Systematic Review Process. In *EASE*. ACM, June 2016. (cited on Page 27, 29, and 31)

Weiguo Fan, Linda Wallace, Stephanie Rich, and Zhongju Zhang. Tapping the Power of Text Mining. *Magazine; Communications of the ACM - Privacy and security in highly dynamic systems*, 49(9):76–82, 2006. (cited on Page 14, 15, 16, and 17)

Usama Fayyad, Gregory Piatetsky-shapiro, and Padhraic Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54, 1996. (cited on Page ix and 14)

Katia R. Felizardo, Norsaremah Salleh, Rafael M. Martins, Emília Mendes, Stephen G. MacDonell, and José C. Maldonado. Using Visual Text Mining to Support the Study Selection Activity in Systematic Literature Reviews. pages 26–35. IEEE, 2011. (cited on Page ix, 1, 2, 16, 27, 29, 31, 32, and 33)

Katia R. Felizardo, Stephen G. MacDonell, and José C. Maldonado. A Systematic Mapping on the use of Visual Data Mining to Support the Conduct of Systematic Literature Reviews. *Journal of Software*, 7(2):450–461, 2012. (cited on Page 14)

Evgeniy Gabrilovich and Shaul Markovitch. Feature Generation for Text Categorization Using World Knowledge. In *19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1048–1053. ACM, 2005. (cited on Page 19)

John A. Goldsmith, Derrick Higgins, and Svetlana Soglasnova. *Automatic Language-Specific Stemming in Information Retrieval*, volume 2069. Springer, Berlin, Heidelberg, 2001. (cited on Page 17)

Vishal Gupta and Gurpreet S. Lehal. A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, 1:60–76, 2009. (cited on Page 13, 15, and 17)

Demetris Hadjitofallis. Automatic Term Recognition. Project background report, The University of Manchester, 2010. (cited on Page 15)

Edgar Hassler, Jeffrey C. Carver, David Hale, and Ahmed Al-Zubidy. Identification of SLR tool needs - results of a community workshop. *Information and Software Technology*, 70:122–129, 2016. (cited on Page 1, 2, and 61)

Marti A. Hearst. Untangling text data mining. In *37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*. ACL, 1999. (cited on Page 13)

Daniel A. Keim. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002. (cited on Page 13 and 16)

Barbara A. Kitchenham. Procedures for performing systematic reviews. Technical Report TR/SE-0401 and 040001 1T.1, Keele University and National ICT Australia, 2004. (cited on Page 5 and 21)

Barbara A. Kitchenham and Pearl Brereton. A systematic review of systematic review process research in software engineering. *Information and Software Technology*, 55: 2049–2075, 2013. (cited on Page 14)

Barbara A. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, Keele University, 2007. (cited on Page ix, 1, 2, 5, 6, 7, 13, 21, 23, 24, 25, 37, 39, 47, and 61)

Barbara A. Kitchenham, Tore Dybå, and Magne Jørgensen. Evidence-based Software Engineering. In *International Conference on Software Engineering (ICSE)*. IEEE, 2004. (cited on Page 1)

Barbara A. Kitchenham, Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. Systematic literature Reviews in Software Engineering - A systematic literature review. *Information and Software Technology*, 51(1):7–15, 2009. (cited on Page 1 and 28)

Barbara A. Kitchenham, Rialette Pretorius, David Budgen, Pearl Brereton, Mark Turner, Mahmood Niazi, and Stephen Linkman. Systematic literature reviews in software engineering – A tertiary study. *Information and Software Technology*, 52: 792–805, 2010. (cited on Page 49)

Barbara A. Kitchenham, David Budgen, and Pearl Brereton. *Evidence-based Software Engineering and Systematic Reviews*, volume 4. CRC Press, 2015. (cited on Page 21, 24, 35, and 52)

Michael Krauthammer and Goran Nenadic. Term identification in the biomedical literature. *Journal of Biomedical Informatics*, 37:512–526, 2004. (cited on Page 15)

Jan H. Kroeze, Machdel C. Matthee, and Theo J. D. Bothma. Differentiating data- and text-mining terminology. pages 93–101. ACM, 2003. (cited on Page 13)

Pavel Laskov and Nedim Šrndić. Static detection of malicious JavaScript-bearing PDF documents. In *27th Annual Computer Security Applications Conference (ACSAC)*, pages 373–382. ACM, 2011. (cited on Page 44)

Christian Lausberger. Konzeption von Suchprozessen und Suchstrategien für systematische Literatur Reviews. Masters's thesis, Otto-von-Guericke-Universität Magdeburg, Germany, January 2017. (cited on Page ix and 42)

Mark MacGillivray and CC-BY Jim Pitman. How to do BibJSON. Website. Available online at http://okfnlabs.org/bibjson/; visited on September 10th, 2017. (cited on Page 38)

Viviane Malheiros, Erika Höhnr, Roberto Pinho, Manoel Mendonca, and José C. Maldonado. A Visual Text Mining approach for Systematic Reviews. pages 245–254. Empirical Software Engineering and Measurement, 2007. (cited on Page 16, 27, 29, and 31)

Inderjeet Mani. *Automatic Summarization*. John Benjamins Publishing Company, Amsterdam, 2001. (cited on Page 16)

Jefferson S. Molleri and Fabiane B. V. Benitti. Automated approaches to support secondary study processes: a systematic review. In *International Conference on Software Engineering and Knowledge Engineering*, pages 143–147. Semantic Scholar, 2012. (cited on Page 2)

Raymond J. Mooney and Un Y. Nahm. Text Mining with Information Extraction. pages 141–160. Multilingualism and Electronic Language Management, 2005. (cited on Page 13)

Srinivas Nidhra, Muralidhar Yanamadala, Wasif Afzal, and Richard Torkar. Knowledge transfer challenges and mitigation strategies in global software development—A systematic literature review and industrial validation. *International Journal of Information Management*, 33:333–355, 2013. (cited on Page 52)

Fábio R. Octaviano, Katia R. Felizardo, José C. Maldonado, and Sandra C. P. F. Fabbri. Semi-automatic selection of primary studies in systematic literature reviews: is it reasonable? *Empirical Software Engineering (ESE)*, 20(6):1898–1917, 2015. (cited on Page ix, 27, 29, 31, and 34)

Jakub Piskorski and Roman Yangarber. *Multi-source, Multilingual Information Extraction and Summarization*, chapter Information Extraction: Past, Present and Future, pages 23–49. Springer Berlin Heidelberg, 2013. (cited on Page 15)

Antoine Ponsard, Francisco Escalona, and Tamara Munzner. PaperQuest: A Visualization Tool to Support Literature Review. In *Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA)*. ACM, 2016. (cited on Page 27, 29, and 31)

Dragomir R. Radev and Kathleen R. McKeown. Generating natural language summaries from multiple on-line sources. *Journal Computational Linguistics - Special issue on natural language generation*, 24:470–300, 1998. (cited on Page 17)

Ivonne Schröter, Jacob Krüger, Philipp Ludwig, and Thomas Leich. Identifying Innovative Documents: Quo vadis? In *International Conference on Enterprise Information Systems*, pages 653–658. ScitePress, 2017a. (cited on Page 2)

Ivonne Schröter, Jacob Krüger, Janet Siegmund, and Thomas Leich. Comprehending Studies on Program Comprehension. In *25th International Conference on Program Comprehension (ICPC)*. IEEE/ACM, 2017b. (cited on Page 12)

James Thomas, John McNaught, and Sophia Ananiadou. Applications of text mining within systematic reviews. *Research Synthesis Methods*, 2(1):1–14, 2011. (cited on Page 13, 15, 16, and 17)

Federico Tomassetti, Giuseppe Rizzo, Antonio Vetro, Luca Ardito, Marco Torchiano, and Maurizio Morisio. Linked Data approach for selection process automation in Systematic Reviews. In *EASE*. IET, April 2011. (cited on Page 27, 29, and 31)

June M. Verner, Pearl Brereton, Barbara A. Kitchenham, Mark Turner, and Mahmood Niazi. Systematic Literature Reviews in Global Software Development: A Tertiary Study. In *International Conference on Evaluation and Assessment in Software Engineering (EASE)*, pages 2–11. IET, 2012. (cited on Page 1)

W3.CSS W3Schools. JSON - Introduction. Website. Available online at https://www.w3schools.com/js/js_json_intro.asp; visited on September 20th, 2017. (cited on Page 40)

Jane Webster and Richard T. Watson. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *Management Information System Quarterly*, 26(2): xiii–xxiii, 2002. (cited on Page 1)

Sholom M. Weissi, Nitin Indurkhya, and Tong Zhang. *Fundamentals of Predictive Text Mining*, chapter From Textual Information to Numerical Vectors, pages 13–38. Springer, London, 2010. (cited on Page 17)

Bas Wenneker. BibMix: Enrichment of Citation Metadata Based on Integration of Bibliographic Data. Masters's thesis, Delft University of Technology, The Netherlands, September 2010. (cited on Page 38)

Ian H. Witten. Text mining. In *Digital Library, International Journal on Digital Libraries archive*. CiteSeer, 2011. (cited on Page 13, 14, and 17)

Ian H. Witten, Eibe Frank, and Mark A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 2011. (cited on Page 13)

Yiming Yang and Jan O. Pedersen. A comparative study on feature selection in text categorization. In *International Conference on Machine Learning*, pages 412–420, 1997. (cited on Page 15)

Mansooreh Zahedi, Mojtaba Shahin, and Muhammad A. Babar. A systematic review of knowledge sharing challenges and practices in global software development. *International Journal of Information Management*, 36:995–1019, 2016. (cited on Page ix, xi, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 59, and 61)

He Zhang and Muhammad A. Babar. Systematic Reviews in Software Engineering: An empirical investigation. *Information and Software Technology*, 55:1341–1354, 2013. (cited on Page 2)

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Magdeburg, den 23.10.2017