

Received March 7, 2022, accepted March 24, 2022, date of publication April 5, 2022, date of current version April 13, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3165079

The Application Domains of Systematic Mapping Studies: A Mapping Study of the First Decade of Practice With the Method

ERNO VANHALA¹, JUSSI KASURINEN¹, ANTTI KNUTAS¹, AND ANTTI HERALA²

¹Department of Software Engineering, LUT University, FI-53851 Lappeenranta, Finland

²MP Soft, 04410 Järvenpää, Finland

Corresponding author: Erno Vanhala (erno.vanhala@lut.fi)

ABSTRACT The systematic mapping study (SMS) is a relatively new method of generating new information from existing studies. First defined as a methodology in 2007, it offers a method to filter existing information to produce novel insight into the observed research domain, and pinpoint new directions of research. In this study, the systematic mapping study method was utilized to determine how SMS as a method has spread and was utilized during the first decade since its conceptualization. In general, it was found that the SMS method is still at its early phase in utilization, and is mainly used in software engineering and healthcare studies, but also in several other scientific domains. SMS research and the scientific outputs rely on transparent protocols when conducting the actual search and identification process, and so far, the applied protocol and research procedure correlates strongly with the application domain; different domains have their own protocols. The SMS method can be recommended, for example, when the aim is to gain knowledge on how a specific topic is studied and where there are research gaps. There are still areas that are debated or where successful implementation is difficult, the biggest problems being the amount of work it requires and possible lack of quality analysis of the articles.

INDEX TERMS Meta study, research process, systematic mapping study, systematic scoping review, tertiary research.

I. INTRODUCTION

In the time before digital access to global research databases, the most common approach to gain extensive knowledge on an entire field of science was by examining a limited number of publications, such as essays, books, theorems, or periodicals. However, in the modern era of science, this is no longer true due to the number and availability of different publications and publication venues. In fact, the more common problem is not that the prior research work is not available or is very limited, but that there is simply too much information to actually form any meaningful views on the current trends or core topics of that field at a glance [1]. The issue of the gargantuan amount of relevant information is problematic, especially in the fields that rely on the evidence-based research, such as medicine, but also applies in other areas, such as software engineering, where the current trends can heavily influence the focus of the research work [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Adnan Abid.

To address this issue, several strategies on how to identify and collect the relevant information [3], [4] have been developed, such as systematic mapping studies (SMS) or systematic literature reviews (SLR), both of which enable the researcher to gain an insight into the entire research domain. The objectives of these research approaches are somewhat different [1], [5]: when systematic literature review focuses on gaining as much precise information and focused in-depth data as possible, the systematic mapping studies tend to focus on the research trends and activity in the field [6]. It is also noteworthy that the approaches have been blended over the years, and that the guidelines are also somewhat flexible in different domains, sometimes even blurring the line between mapping study and literature review [7]. In any case, the objective of these methods is to enable the researchers to gain reliable insight into the current trends and topics of interest, and identify potential research gaps in the topic at hand. Interestingly, due to access to digital databases, these methods also seem extremely flexible on the applicability between the different fields of science and scientific research.

Based on these distinctions, our research group decided to conduct a study on the applicability and spread of the SMS methodology in the different scientific domains after its introduction in the mid-2000s. The focus of our research was on the applicability of the systematic mapping studies in the different fields of science, and to examine how the SMS method has been adopted in academia during its introductory decade of applications. To assess the usability and level of utilization of the method, our group conducted a systematic mapping study of systematic mapping studies to gain insight into the method itself and collect further information on the applicability and applications of the method in practice. The following research questions were formulated to guide this research:

- RQ1: Which research fields utilize systematic mapping studies?
- RQ2: Where and when were the mapping studies published?
- RQ3: Which guidelines – or protocols – are used when conducting systematic mapping studies?

The analyzed systematic mapping studies were identified from a group of 254 different scientific domains from acoustics to zoology, as defined by the Web of Science research domain taxonomy [8]. Overall, our study analyzed 423 examples of applied systematic mapping studies during the first decade the method was actively applied. These publications were further analyzed with both qualitative and quantitative approaches to understand how and where the SMS method has been adopted. Although there exist a few systematic mapping studies on the systematic mapping studies (e.g., [9], [10]), they are all meta-analysis on a specific field of research. In this study, the objective is to conduct a comprehensive systematic mapping study of the systematic mapping study method itself, taking into consideration all scientific domains and application methods between the years 2007 and 2017, the first identified year of application and the decade following it.

II. BACKGROUND AND RELATED RESEARCH

To begin this research, it is relevant to understand the difference between the concepts of a systematic literature review (SLR) and a systematic mapping study (SMS). Although on some occasions they can produce overlapping results, the objectives between these concepts are different. The SLR aims to generate precise deep information on a focused topic [1], [11]. Collecting this much precise information is more resource-intensive, which would mean that when properly applied, SLR should not be conducted on topics with a wide scope. Instead, the SLR method excels when the research questions are kept specific, the number of publications is low, and the material contains only minimal amounts of non-relevant information. In fact, Kitchenham *et al.* [3] argue that in SLR, all relevant studies should be found since the protocol aims for in-depth analysis of the given topic.

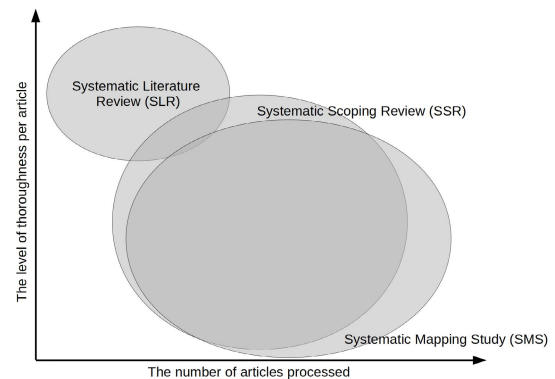


FIGURE 1. The positioning of systematic literature review, systematic scoping review and systematic mapping study in the axis of article broadness and analysis depth. Based on [1], [5], [11], [15], [16].

The systematic mapping study, on the other hand, can have multiple research questions with a broader scope than SLR, as the focus can be directed towards a wider scope, such as a research trend and activities on a field of research [1], [11]. The SMS aims to provide an indication of the quantity of the evidence, whereas SLR aims to provide an indication of quality [11]. Because of this, in systematic mappings it is not required to find every single piece of research evidence on a topic, but rather a representative sample of the relevant studies [3], [12]. Nevertheless, SMS is still a review, but the aim is not to discuss the findings of an individual article, but to create a general overview and to define the big picture of the search domain [5]. In the end, the aim of a systematic mapping study is to draw a map, which is a type of conclusion of found ideas, research gaps, directions of research, or any other concise representation of the research. The map can be represented by, for example, a table, a diagram or, a flow chart.

In some cases, the number of articles found in the systematic mapping study is low enough to allow researchers to extend the SMS with parts of SLR or even do both to some extent (e.g., [13], [14]). In this kind of scenario, the collected metadata and identified research gaps can be extended with a deeper analysis of topics discovered during the research process. Similarly, there are also systematic scoping studies, which are described as a synonym for or at least very closely resembling systematic mapping studies [5], [11], [15]. In general, scoping studies also share some characteristics with systematic literature reviews, which are missing from the systematic mapping studies, meaning that they tend to include aspects from both SMS and SLR methodologies, or are SLR-supplemented SMS studies. Fig. 1 illustrates these three concepts.

Table 1 concisely represents the differences between the systematic literature review and the systematic mapping study / systematic scoping review methods. This representation is only a rough categorization since these methods can be used in parallel and there is not necessarily a clear distinction

TABLE 1. Stereotypical comparison of systematic literature review and systematic mapping study, based on [1], [5], [11], [17].

	Systematic literature review	Systematic mapping study / systematic scoping review
Aim	To gain focused and deep knowledge	To gain broad coverage and data for classification; to find research gaps
Number of processed articles	Low	High
Indication sought	Quality	Quantity
Analysis of articles	Thorough read-through, meta-analysis	Keywords, summaries, graphical representations, key statistics
Focus	A single, specific topic	Broader scope
Results	Deeper knowledge on the specific topic through qualitative explanations	Mostly a quantitative representation of the field or topic

on which method to apply in a given scope of research. In a sense, the method is defined more by the results than the applied steps.

The guidelines for the SMS are also a varying set of different methodologies. The two commonly applied methods from Petersen *et al.* [16] and Kitchenham and Charters [11] further illustrate this point and shared nature of the SLR and SMS methodologies; Petersen *et al.* discusses and defines the SMS approach in great detail, whereas Kitchenham and Charters define the SLR studies, with SMS only as a phase or step towards systematic literature review. Other common SMS methodology definitions such as [18] or [19] focus on defining the systematic mapping studies and their application in scientific domains other than computer science or systems engineering.

III. RESEARCH PROCESS

One of the first steps in the systematic mapping study is the identification of the appropriate venues and sources from which to collect the data material [16]. After initial trial searches with several domain-specific search engines and databases, such as IEEE Xplore and ACM digital library, it was decided to focus on the domain-neutral Google Scholar as the main source of data. The utilization of Google Scholar was considered to yield less domain-biased results than utilizing domain-specific libraries – although this tool cannot access every existing scientific database or publication venue, it should offer a representative set of the overall application of the protocols and the related research fields. The application of more domain-specific libraries such as ScienceDirect, PubMed, or IEEE Xplore would have biased the dataset towards certain scientific disciplines, whereas using one generalist search engine as the data source gave equal coverage to all scientific venues and research domains. Finally, it has also been argued that Google Scholar is a suitable tool due to its convenience, low cost, and broad coverage [20]. Gius-tini and Boulos [21] argue that the Google Scholar cannot be applied in the systematic reviews because it finds only

TABLE 2. The systematic mapping study process based on the model presented in [1].

Step #	Process Activity	Output
1	Defining the research question, identifying the relevant forums	The scope of the mapping study
2	Conducting search for the primary studies in the topic	The paper pool of the systematic mapping study (all papers)
3	Screening and identifying the primary papers	The relevant papers from the pool of all papers
4	Keywording using the abstracts	Classification of the papers
5	Data extraction and mapping process	Systematic map, statistical data
6	Documentation of the systematic map, analysis of the other results	Systematic mapping study

approximately 95% of the articles in the controlled test group, and includes a large number of irrelevant objects. For our purposes, we consider 95% accuracy reasonable, since our aim was to gather articles from as wide a range as possible instead of precise accuracy. A brief comparison of Google Scholar to ACM Digital Library, IEEE Xplore and Scopus showed how Google Scholar found 13.1 (ACM), 4.9 (IEEE) and 1.5 (Scopus) times more results with the same search settings. The systematic mapping study approach applies different phases from data collection to the results to refine the data towards documentable results. In our mapping study, we applied the process model as defined in Petersen *et al.* [1], which has six process steps; definition of the scope, search for the paper pool, identification of primary documents, classification, data extraction, and documentation. These process activities are summarized in Table 2.

Besides Petersen *et al.* [1], principles presented by Grant and Booth [22] and Kitchenham *et al.* [17] were used in the design of the study. In the following sections, each process step is explained in practice.

The research question definition was accomplished by establishing where the systematic mapping studies were discussed and published prior to this publication, and studying whether the systematic mapping study metadata had been comprehensively analyzed. This step was conducted by the first author in order to identify whether the research questions for this work would produce novel information or propose enhancements to the research method. Once established that the objectives could yield novel information, the scope was set and the research questions were formulated. Along with the research question formulation, the first iteration of the data collection tools were designed to systematically collect pieces of relevant information. It has been suggested that the research questions for systematic mapping studies should be formulated in parallel with the design of the classification framework, since it is the tool that is used for the mapping of the individual publications in the planning stages of the research process and dictates which data are produced as output for further analysis [23].

The search for primary studies was conducted on Google Scholar by formulating a list of rules and keywords. The

TABLE 3. Keywords used in various searches and the number of results they provided.

Search round	Search phrase (in title only)	Results in Google Scholar
R1	"systematic mapping study"	405
R2	"mapping review"	130
R3	"systematic mapping studies"	32
R4	systematic "scoping review"	191
Total		758

initial search keywords were 'systematic mapping study' in the title of an article (Table 3, R1). This was a logical starting point as the goal was trying to gain knowledge on systematic mapping studies. To keep the number of irrelevant or non-scientific documents as low as possible, it was required that the keywords existed in the title of an article. Adding the name of the applied method seemed to be a de facto standard in scientific publications when naming articles reviewing the literature. When searched from anywhere within the article the number of search results increases tenfold, resulting in a large set of irrelevant articles, which would have required an unfeasible number of resources to analyze everything at a satisfactory level.

An initial search showed that approximately 80% of the studies were relevant. This was considered a success and it was decided to continue searching for the title of the article only and keep the keywords in between quotation marks. A quick review of the found articles also mentioned that the mapping studies were also described as 'mapping reviews' [5]. These findings led to the second search round.

The search string 'systematic mapping review' initially provided only 32 hits and it was decided to try the search without the keyword systematic. This search (R2) also yielded relevant articles, although it also included more irrelevant articles. It was decided to search also for 'mapping study' without the keyword systematic. This search backfired and resulted in mainly irrelevant findings as the keyword mapping was considered to be related to cartography. It was also decided to try the plural 'studies', as it seems that some articles were discussing more than one mapping study. This search round (R3) produced only a handful of articles.

The final search round (R4) was done with the keywords systematic 'scoping review' as it was noticed that health-care and medical science research used this term instead of 'mapping study' while conducting fundamentally similar studies. This search round provided almost only healthcare- and medicine-related studies since the keywords had strong domain-related connotations. While systematic mapping studies and systematic scoping studies appeared synonymous, the search engine does not understand this concept, meaning that these method names had to be searched separately, resulting in different numbers of results.

The searches were conducted in June 2020, and were not restricted by any rules or filters except the ones mentioned earlier in this section and the year limit of up to 2017. In practice, at this stage, all of the articles, regardless of

TABLE 4. Key figures of the search rounds and their findings.

Search round	Results in Google Scholar	Achieved additional articles	Accepted articles	Acceptance percentage
R1	405	347	259	75%
R2	130	104	58	56%
R3	32	16	9	56%
R4	191	134	97	72%
Total	758	601	423	71%

the publication year, publication forum, language, or actual content were included.

In Table 4, key figures of search rounds are presented. 'Achieved additional articles' indicates the number of new articles found as an extension to previous search rounds – thus duplicates are removed. In round one, not all articles could be included due to copyright issues, not being publicly available, or being non-scientific, such as presentations. These were also removed from later search rounds. Acceptance percentage is calculated from the accepted articles divided by the achieved additional articles.

A manual search was not conducted beyond pilot studies with keyword feasibility, nor did we perform snowball sampling. This was found to be justified since it was found that database search is the most used, followed by manual search and snowball sampling with far less use [1]. In this case, the manual search might have biased the study, as the goal was to gain articles from all areas, but having prior experiences primarily from the different computer science and software engineering domains would have introduced bias. The problem with snowballing would have been the number of articles gained from snowballing. The study by Petersen *et al.* [1] alone, which was referenced in several of our studies, has more than one thousand citations according to Google Scholar (at the beginning of 2021), and other mapping study protocols had a citation count even higher than that. It would have required immense human resources to go through all these articles in the required detail, and a short pilot search of the snowballing data indicated that the number of irrelevant citations increases rapidly. With these points in mind, it was decided to concentrate on conducting an accurate database search to gather as wide and inclusive a set of articles as possible.

The research group conducted the screening and identification process manually. In this step, the objective was to identify the primary documents from the pool of all documents generated by the search process. As defined in [16] and [24], the applied inclusion criteria principles were as follows. If the abstract explicitly mentioned 'systematic mapping study' or a similar term, the paper was included into the primary documents. Only in the cases where the appropriate terms appeared in the keywords or index terms but were not mentioned in the abstract, or the abstract clearly indicated that the research work did not include systematic mapping study, was the paper excluded. In this step, some

sanity checks were applied, such as removal of the duplicate entries and removal of partial, broken, or non-research papers. Although most of the cases were straightforward to include or exclude, the researchers of the group stayed in constant communication throughout the screening process and when inclusion or exclusion issues arose, a group consensus was formed. The screening was conducted by authors 1, 2, and 4 while Author 3 concentrated on correlation calculations and computational outputs.

Keyword extraction was conducted by all of the authors. To maintain internal validity, the keywords were extracted by the authors manually to minimize the number of misclassifications. The extraction was done in two steps. In the first step, all of the papers were classified by their content and context. After the initial set of keywords was given, they were grouped into larger abstract terms to establish categorizations for the primary papers and provide initial clustering to classify papers into groups. When the final sets of keywords and categories were defined, the documents could be clustered and the mapping process could be started. However, it should be noted that, in practice, the definition of the keywords and categories are parallel activities; new articles define new keywords, which create new categories, merge into existing categories, or divide existing categories into sub-categories. Because of the relatively low number of articles selected for this part of the study, it was safer and more feasible to follow the manual practice in classification instead of introducing more error-prone automatic procedures.

To minimize the differences between different authors conducting the classification process, we applied a classification scheme and data extraction form to minimize the noise and ambiguity generated by having several people working on the data at the same time. The classification scheme in this work followed the concepts presented by [25] since the research scope in this work included all possible systematic mapping studies, not only a subsection or certain domain. However, in addition to the types of research papers, our classification also collected types of research domains, based on the classification by Web of Science [8].

Fleiss' kappa [26] was run to determine inter-rater reliability in data extraction. A random sampling approach (CL.9) was used to select evaluated data and raters from the dataset. Fleiss' kappa showed that there was a good agreement between the raters at $\kappa = .704$ (95% CI.701 to.708; $p < .0005$).

The data extraction and mapping process consisted of several different methods and mathematical models. The systematic mapping of the primary documents and their keywords were compiled for the visual presentation of our work. Additionally, structural information such as the applied protocols, number of accepted primary documents, and topic-independent data classifications were identified from the primary documents during the data extraction process. The complete list of collected data items is listed in Table 5.

The quality of the articles and the merits of the research work itself were not evaluated in this study. As suggested by

TABLE 5. The data items recorded to the data collection instrument.

Recorded data item	Description
Work title	The title of the document
Authors	The authors of the document
Publication year	The publication year of the document
Country of origin	The countries the authors have provided as their contact information. Some studies had researchers from several countries; all were noted.
Type of document	What type of contribution the document is, possibilities being conference proceedings paper, journal article, workshop paper, thesis study, or other.
Peer reviewed document	Is the document from a peer-reviewed source? Yes/No
Origin of the document	Are the authors from an academic or industry research group?
Field of science	What fields of science the document represents, as based on [8]
Type of systematic mapping study	What type of SMS contribution the document is? Study/Guideline/Both/Other
SMS or SLR	Is the research work presented in the paper actually a systematic mapping study, systematic literature review, both, or something completely different?
Primary guidelines or protocol applied in the article	What protocol was used to generate the steps that were utilized in the study. Some studies were based on more than one protocol; all were noted.
Search strategy and data source type	What methods and sources were used in the search for the documents
Development of the search string	What method was used to generate the keywords and the search string
Inclusion and exclusion process	What method was applied to the document inclusion and exclusion process.
The initial pool of documents	The number of documents at the starting point from which the researchers started to narrow down to the most relevant core studies
Number of primary documents	The number of documents finally accepted to the systematic mapping study and analysis
Topic-independent classifications	What topic-independent elements are listed and described (e.g., histogram of years of publication)
Quality assessment methods applied	What methods have been applied to make sure that the accepted publications are the ones that matter, and have the acceptable quality to be considered for inclusion in the study
Research questions	The list of research questions – or study aims if no research questions were formulated
The format of the data presentation	What kind of illustrations were utilized to present the findings of a study
Number of references	How many articles cite this study, as identified by Google Scholar

other researchers (e.g., [1], [11]), the quality of articles should not be considered as a major concern in systematic mapping studies in general. Mapping studies aim for a wider number of articles than systematic literature reviews, and since this mapping study collected information on all scientific domains, the decision was made to not include quality analysis beyond the point, and that some form of mapping study approach was applied. Besides, the different scientific domains have different conventions and priorities in their publication styles, which would make the quality evaluation of individual articles difficult, prone to errors, and in the worst case, steer

the results towards certain scientific domains. In the end the exclusion criteria were the following:

- 1) The non-peer reviewed articles were rejected
- 2) The works actually conducting systematic literature review instead of mapping study, even though they reported the latter were also rejected.
- 3) The articles that were not written in English, Spanish or Portuguese were rejected as we could not aggregate data from other languages
- 4) The work has been published during the period of 2007 to 2017

Since the earliest identified SMS study accepted to the primary data set was from the year 2007, this made the study examine the first decade of SMS research. With this time span it was possible to understand the beginning of the utilization of the method. We are documenting the first generation of protocols and their spreading over the scientific fields. Documentation of the work and analysis of the statistical data included activities and steps needed to develop the results presented in the next chapter. The systematic mapping process with the map was generated following the principles presented in [16] supplemented with the concepts presented in [1], whereas the actual statistical data analysis was conducted with R, a statistical analysis language and toolset [27]. Three descriptive methods were used to quantitatively analyze and present the data:

- 1) Basic descriptive statistics of sums and means
- 2) Co-occurrence matrices and tetrachoric correlations to discover similarities between publications
- 3) Topic modelling to statistically sort publications into groups

Python scripts were used to collect, collate, and sum references. Python standard library-based difflib measure of string distance – the similarity between inputs – was used to correct different ways of spelling and to consolidate data. Visualizations were generated with a spreadsheet program. In the correlation analysis of the item feature co-occurrence, the polychoric R library [28] was used to calculate tetrachoric correlations, which is a well-suited method for dichotomous datasets [29]. The p-value was adjusted with the Holm-Bonferroni method [30] for multiple comparisons to control the family-wise error rate.

The documents were also sorted into topics using the latent Dirichlet allocation (LDA) algorithm [31] using a modified version of the NAILS script [32], which utilizes the topicmodels R package [33], and visualized with the LDavis library [34]. LDA can be used as a statistical method text mining method for assigning documents into topics, which are detected using word association and distributions [35]. The underlying mechanism in LDA is a probabilistic Bayesian network model in which each document is characterized by certain topics, and each topic is defined by a specific set of words, which co-occur with a certain probability. To summarize, the topics of each document are defined by a set of words that often appear together. Semantic coherence, a quality

value for deciding the number of topic models [36], was calculated using the R stm library [37]. Additionally, the LDavis library was also used to calculate the distance between topics on a scatterplot, which approximates the semantic relationships between the topics with multidimensional scaling. It is a method similar to factor analysis and allows the level of similarity between objects to be visualized. The inter-topic distance was calculated using Jensen-Shannon divergence [34]. LDA-based topic modelling is a commonly used method for text analysis and equivalent methods have been used to statistically analyze scientific texts in previous studies [38]–[41].

IV. RESULTS

In this section, the results of the systematic mapping of the works applying systematic mapping study as their primary research method are presented. First, general statistical information is presented to establish fundamental observations regarding the research method, followed by more in-depth analysis with the systematic mapping of the different domains and observations.

A. BASIC INFORMATION AND STATISTICS

Based on our analysis of the different types of systematic studies, our research team identified 601 different studies and documents describing systematic studies. After classification and assessment, we accepted 423 documents as our pool of primary studies, since they included complete research conducted with SMS methodology, were published during the period of 2007-2017, and were not a duplicate or earlier versions of other studies. Out of these documents, roughly two thirds were journal articles, and the rest were different types of conference or workshop publications. Out of the primary documents, only 13 studies (3.1%) were from industrial sources, where at least one of the authors was from outside academia. Overall, 1,466 individual authors were identified from the dataset of primary documents; on average 4.3 authors per publication. Only eight publications (2%) were published by just one author, which can be considered as a good sign related to certain inherent risks such as the researcher bias (see for example [22]), but also as an indication of the general effort required by the application of the systematic mapping study methods. Additionally, out of 1,466, 164 authors had published more than one systematic mapping study, with 39 authors being involved in three or more.

Of the primary documents, the systematic mapping studies achieved the initial, median dataset of 1,278 documents, which during the classification and categorization converts into a median pool of 58 primary documents. The highest number of citations was collected by a study published in 2008 in the topic of software engineering, and it had at the moment of data collection (in 2020) more than 2,700 citations. Table 6 summarizes quantitative metrics of the included primary articles.

TABLE 6. Basic figures on the examined systematic mapping studies.

Item	#	Type
Number of journal articles	263	documents
Number of conference proceedings full papers	148	documents
Number of identified authors	1466	authors
Number of authors with more than one systematic mapping study published	164	authors
Studies with the majority of the authors from non-academic organizations	13	studies
The median size of the initial dataset in systematic mapping studies:	1278	documents
The median size of the primary document dataset:	58	documents

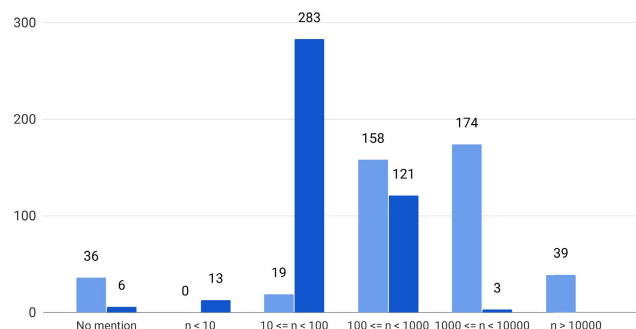
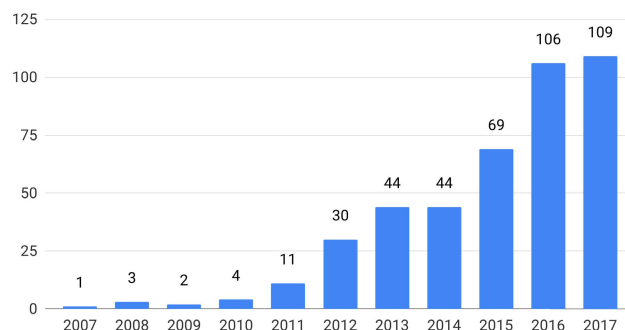
**FIGURE 2.** The x-axis represents the number of selected articles (n), where the light bars illustrate the set of initially found publications and dark bars illustrate the set of final, accepted, publications. The y-axis represents the number of articles selected for this study in each category. Three articles had two different studies included in it.

Fig. 2 describes how the number of publications is reduced from the initial search results to the set of accepted articles. In 39 studies, the initial number of articles was over ten thousand, with 174 studies having one thousand or more and less than ten thousand articles. In the final accepted set of primary publications, three articles had more than one thousand studies included. Nineteen studies reported having an initial number of publications of less than one hundred whereas the final dataset was less than one hundred in 296 cases (three articles had two different studies combined) and even less than ten in 13 cases. In a few studies, there are no mentions of the number of articles and it was not manually calculated from the references.

As the aim of the study was to improve the understanding of how the systematic mapping studies are used and how they were adopted by the academic community, it is given that the publication volumes in the initial years would be low. In general, the trend shows healthy growth during the observed period; the number of published studies increased from fewer than ten publications per year (<2011) to more than 100 per year since 2016. This tenfold increase indicates that the method has gained popularity among the researchers, settling to the level of more than one hundred per year in 2016 and 2017. Fig. 3 illustrates the growth trend of the systematic mapping studies.

The origin of the studies also indicates how the systematic mapping study method has gained a foothold in academia.

**FIGURE 3.** The number of published systematic mapping studies has increased during the second decade of this millennium.**TABLE 7.** The countries most involved in the published systematic mapping studies, with the comparison against the total science output rating [42].

Country	# of SMS studies	Position shift against SCImago output rating
UK	76	+2
Brazil	73	+13
Canada	50	+6
Spain	37	+6
USA	31	-4
Sweden	26	+3
Australia	24	+4
Finland	16	+19
Germany	16	-5
Malaysia	14	+19
Morocco	11	+46
Pakistan	11	+34
Turkey	11	+9
Italy	10	-4

Overall, it was discovered that researchers from more than 50 different countries or self-governing areas had applied the method in at least one published and peer-reviewed scientific publication during the observed period. Based on the publications, the Brazilian researchers were very active early adopters in the production of the systematic mapping studies, along with people located in the UK. Other early adopters who had large contributions to the SMS studies in relation to their relative average contribution to the scientific publications were Finland, Malaysia, Morocco, Pakistan, and Turkey. Table 7 illustrates the countries where the method usage was most popular during the period of 2007 to 2017, and their systematic mapping study ranking against their current overall scientific output ranking in August 2020 [42].

Where researchers have been utilizing the systematic mapping study method globally, the same cannot be argued about the field and domain distribution. Web of Science Research Areas were utilized as a tool to categorize all of the collected studies [8]. The Web of Science categorization includes 254 different categories in a wide range from acoustics to zoology, and because of its wide scope, it was considered to be suitable for this study. By using the Web of Science categorization, we found out that 82.2% of the collected studies have been published in three fields: software engineering, health care, and medicine. This was a major drawback when

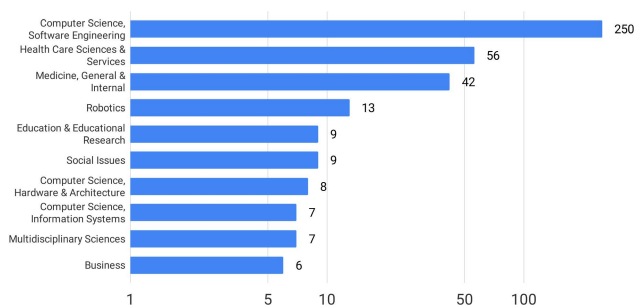


FIGURE 4. Distribution of studies between various fields, x-axis representing the number of studies. Comprehensive count is available in Appendix 2 in Table 12.

arguing how a systematic mapping study could be applied to different topics, or when assessing the global usability of the method in various domains. However, with more detailed analysis it was observed that besides these three leading fields the systematic mapping study has been used in various domains, but mostly sporadically; in domains such as business, telecommunications, education, and political science, systematic mapping study was applied to generate insight into some topics. Overall, different types of systematic mapping studies from 44 different fields of science were identified. Fig. 4 illustrates the ten fields utilizing systematic mapping study method the most, and the rest of the identified categories are listed in Appendix 2.

The publication channels for the systematic studies reflected the distribution of the research domains. The two most common publication venues were related to computer science and information technologies, with the most common venue having 28 publications. These two journals, *Information and Software Technology* and *Journal of Systems and Software* were also the only publication channels that had more than ten SMS studies released. However, due to the naming conventions and limitations of the analysis, it needs to be acknowledged that some of the conference series could possibly reach similar numbers if all separate tracks and workshop proceedings were counted as one publication venue. Overall, over 300 different publication sources were identified from the data, indicating that the systematic mapping study publication venues are very diversified.

Besides observations on the origin countries and research domains, some observations on the application of different strategies and research methods can be observed. On the development of the search string and keywords, only 46 primary studies (10.8%) applied some form of search string development procedure, such as the population-intervention-comparison-outcome (PICO) method defined by [11]. In general, in most cases, the search string did not evolve beyond the first application, as only in 18 cases (4.2%) were there reported any iterative process to develop the search string or keywords involved.

In the identification of relevant primary studies, the application level of strategies such as snowball sampling [43] was

TABLE 8. Basic characteristics of the applied systematic mapping study methodologies in the analyzed documents; N = 423.

Characteristic	N	% (out of 423)
Data presented in non-visualized form (e.g., tables or bullet lists)	327	77.3%
Applies some form of teamwork to ensure classification quality	190	44.9%
Applies some form of manual search activity, besides browsing online databases, to collect documents	84	19.9%
Applies some form of snowballing technique	80	18.9%
Data presented in some form of a map (e.g., bubble map, plot map, actual map)	68	16.1%
Applies some form of data extraction or classification scheme to systematically collect data	64	15.1%
Applies some form of defined keyword development procedure, for example, PICO [11]	46	10.9%
Applies some form of iterative process to develop the keywords and the search string	18	4.3%

somewhat more positive than the application of a defined process for developing keywords or search strings: 80 primary studies (18.9%) involved some form of snowballing in the identification process of the primary documents, or in the assessment of the search accuracy. This is interesting, since for example [1] promotes snowball sampling as a separate, critical step in the systematic mapping study research process. Similarly, manual search activities such as browsing proceeding books, selected journals, or other sources besides online databases was applied in 84 primary studies (19.8%) as a strategy to collect more primary documents or achieve better search accuracy. Data collection schemes and quality assessment models, as defined for example in [44], were applied in 64 primary studies (15.1%). Overall, the most common quality assurance method relied on some form of teamwork or group effort to assess and manage the data coherence; 190 primary studies (44.8%) included these types of activities. For inclusion into the dataset, the most common independent classifiers were related to the year, venue, and paper type. Other aspects, such as citation information, impact factor, or the applied data sources were identified from the data, but their usage in classification or quality assurance work was very uncommon. More details on these aspects are available in Table 8.

To conduct a systematic mapping study, it is a beneficial practice to select a protocol to be followed throughout the study [16]. For example, the protocol by [16] was applied in 156 primary documents (36.8%), making it the most applied protocol in this study. On the other hand, some studies reported no references to any existing protocol, and either utilized some newly developed protocol or lacked the necessary citations.

Additionally, 63 studies were listed more than once as a source for the protocol used in a mapping study. Several studies relied only on one protocol reference, but in some

TABLE 9. Tetrachoric correlations between the field of science and review protocol authors. Holm-adjusted significances marked as * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

	INFORMATION	COMPUTER	HEALTH	SOCIAL	BUSINESS
Budgen	0.06	0.26***	-0.57***	0.00	0.38***
Booth	0.16***	-0.56***	0.31***	0.50***	0.02
Petersen	-0.09	0.67***	-0.68***	-0.44***	0.15**
Kitchenham & Charters	-0.19***	0.66***	-0.67***	-0.29***	0.23***
Arksey	0.04	-0.60***	0.70***	0.23***	-0.09
Levac	0.19***	-0.57***	0.68***	-0.01	0.06
Moher	0.05	-0.61***	0.65***	0.35***	-0.10*
Grant	0.17***	-0.56***	0.30***	0.52***	0.03

cases even four protocols were used when developing the steps conducted in the study. Out of 28 protocol references that were cited more than once, two were considered as major sources for the protocol development. [16] was used as a base for a protocol 158 times and [11] was used on 93 occasions. Both of these studies are aimed for research on software engineering, but they were also cited in studies considering other areas, such as linguistics and business (e.g. [45]–[48]).

The second group of protocol references consisted of [19], [22], [49], [50], [51] and [18]. These sources have been cited more than ten times and were used in areas other than software engineering. For example, a Prisma statement by [18] was widely used in healthcare studies.

On the protocol definitions, it is also worth observation that the one in [11] is a further developed version of the one in [49], yet the latter is still widely used as a base for the protocol. A similar development also happened between [1] and [16]. Finally, there were 80 studies that were used as a basis or a guideline only once or twice. Some of these studies were older than the more applied protocols, so the original concepts might have been presented in papers not covered in the identified protocols, but in this scenario, it is also apparent that they had a very limited impact. Overall, eight different protocol papers (out of these 80) were cited more than ten times, with the combined total of over 400 citations.

In addition to basic numerical or statistical data, a dichotomous co-occurrence matrix was created, using the fields of science and protocol to examine whether there is a difference in the used protocols between different fields. The fields of science were tagged using the Web of Science definitions and the review protocols were tagged by authors. The correlation results are presented in Table 9. According to [52] scale for correlation strength coefficients between 0.4 and 0.6 are moderate. Higher than that, they are strong and lower coefficients denote weak correlations. As we can see from the dataset, the protocols in use are spread among different fields. Computer science and telecommunication fields follow [16] as well as [11] (strong correlation). Health and medicine conversely follow [19], [51], and [18] (moderate to strong). Social sciences are similar to health and medicine and dissimilar to computer science.

TABLE 10. Tetrachoric correlations between the field of science and visualizations. Holm-adjusted significances marked as * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

	INFORMATION	COMPUTER	HEALTH	SOCIAL	BUSINESS
Diagram or chart	0.04	0.15**	-0.43***	-0.10*	-0.08
Table	0.16***	-0.34***	0.24***	0.27***	-0.10*
Figure	0.11*	-0.16***	-0.01	0.02	0.03
Plotmap, model, or scatter plot	-0.08	0.65***	-0.57***	-0.25***	0.07
Graph	0.19***	0.44***	-0.33***	0.03	0.18***

A similar tetrachoric correlation analysis was used to calculate correlations between fields of sciences and the used visualizations. The correlation results are presented in Table 10. As with used protocols, the methods of summarizing the results differ between the fields. Information sciences and business sciences seem to prefer word clouds and other graphs. Health and medical fields strictly prefer tables, which in turn have an inverse correlation to computer science.

B. CLASSIFICATION OF ARTICLES

The LDA topic modeling process divided the papers into five topic categories, based on their content. The number of topics was selected based on semantic coherence analysis. The topics and their distances are visualized in Fig. 5¹ and the ten most common keywords for each topic are listed in Table 11. There are three distinct groupings in topics. The first cluster of topics is related to software engineering (T1) and computer science (T2). The second major cluster are related to healthcare outcomes (T3, T4). Of these topics, T3 is more related to education and social sciences, moving it closer to T1-2, whereas T4 is more concentrated on patient outcomes, interventions, and health. Finally, the third distinct theme (T5) is related to clinical medical trials. Based on this mapping, it appears that the currently dominant themes in systematic mapping studies are computing-related topics, healthcare and education, and medical reviews. However, our findings from manual analysis (Fig. 4) do show that a smaller number of systematic mapping studies also exists in social and business sciences.

V. IMPLICATIONS AND DISCUSSION

In this article, we present the results of a systematic mapping study focusing on the spread and applicability of the systematic mapping study as a research method, and on the application domains, protocols, and practices applied to them. The results were collected from 423 different primary studies applying either systematic mapping study (SMS) or systematic scoping review as their main research method from the first ten years (2007–2017) of the method being introduced and applied by the scientific community. In a number of

¹See interactive visualization at https://aknutas.github.io/sms_on_sms_ldavis

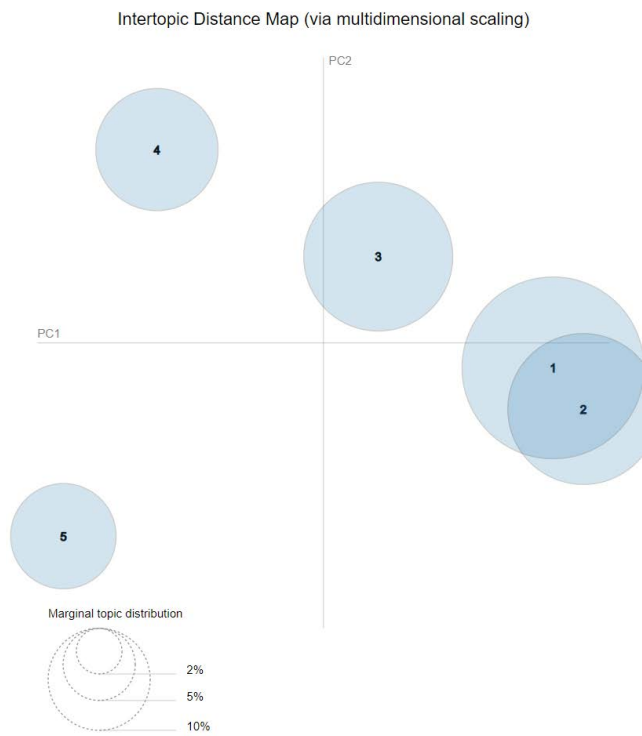


FIGURE 5. Visualizing the LDA topic model and the relative distances in model vocabularies.

TABLE 11. Ten most significant terms describing each discovered LDA topic.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
software	model	health	patient	trial
engineer	test	care	outcome	treatment
develop	computer	education	intervention	group
product	system	quality	care	patient
requirement	application	social	health	risk
model	service	include	include	participant
intern	base	service	age	intervention
search	data	evidence	physical	eczema
system	web	intervention	children	safeties
architecture	cloud	develop	effect	nih

documents, the SMS also delved deeper into the data analysis, incorporating aspects from the systematic literature reviews (SLR).

Overall, the systematic mapping studies are very coherent and process-oriented research works, with a majority of the articles following one or two of the most common identified protocol guidelines. However, there is an argument that there exists a trend of simplification in the application of the different supporting features of the systematic mapping study methods: only a minority of studies apply aspects such as definition process for the keywords (10.9%), iterative search process (4.3%) or data extraction scheme (15.1%). Even the more common methods such as snowball sampling (18.9%) or group work-based quality control (44.9%) exist only in a minority of works, even though they are considered important

by the two most common systematic mapping study research protocol [11], [16]).

There also exists a major trend in the publication popularity for the systematic mapping studies. In this dataset, the median publication year for studies is 2016 when the data range is from 2007 to 2017. This would indicate that the impact of the systematic mapping studies could not be reliably assessed as a whole at this point, since there has not been enough time for the reference pool to gather and application domains to stabilize, so analyzing the number and types of references is not useful. However, even after the initial ten years there already were studies with the citation counts in hundreds, so the argument that the systematic mapping studies, and systematic reviews in general, provide vague results that are not useful outside their inherent ability to identify related research to the authors, can be dismissed since there is clear evidence that the studies are also beneficial to other user groups. This observation is in line with [53], who argues that even though the systematic reviews offer little specific guidance, they map the areas of interest in the context of the studied phenomena.

Another major observation from the fundamental data and measurements was the location of the systematic mapping study works. There is no simple explanation for why some countries are above their general level, besides the United Kingdom and Sweden, which host the home universities for two of the more common SMS research protocols: [11] and [16]. In any case, these results would indicate that the SMS protocols are generally applicable by every expertise level, and do not impose difficult requirements or require expensive specialization to achieve and generate publishable academic results. However, it is interesting to observe that the protocol applications vary from domain to domain; the systematic mapping studies are clearly fragmented to families applying different definitions and approaches, somewhat similarly as, for example, Grounded Theory has divided into Straussian and Glaserian approaches [54]. Fig. 6 illustrates the division between the most applied protocol papers, and the different scientific domains and their protocol preferences. In this figure to minimize the effect of different versions, Petersen and Kitchenham authored protocol papers and their extensions or revisions are combined to groups representing all their works.

One of the aspects in this study was not only to identify the domains that apply the systematic mapping study method but also which areas do not use the approach. Based on the results, it was possible to formulate a search method that identified systematic mapping studies from 55 different areas of research by following the Web of Science classifications [8]. Even though the domains for the application of the SMS methods are dominated by healthcare and computer science, there are studies from other areas such as art, management, business, and humanities in this dataset. In fact, by applying topic modelling and group analysis methods, it was possible to identify five dominant domains: software engineering; computer science; healthcare outcomes related

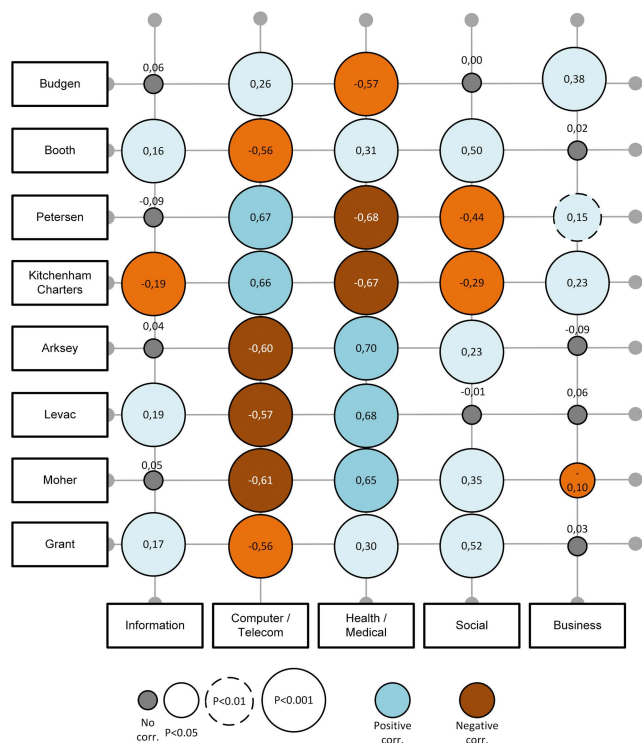


FIGURE 6. Correlation map between the applied SMS research protocol and scientific domains.

to education and social sciences; healthcare outcomes related patient outcomes, interventions, and health; and clinical medical trials. This cross-selection would indicate that the systematic mapping studies are suitable or at least applicable in various domains from humanities to business to science and engineering, even though the methods may not be applied in practice. Additionally, based on this observation, it can be argued that the decision to apply Google Scholar as the multidisciplinary search engine was appropriate, though it should be acknowledged that should the target domain be more specific, as noted by [6], [21], a more domain-specific search engine or database would have been the appropriate search venue.

VI. LIMITATIONS OF THE RESEARCH

As other systematic mapping studies have identified [6], there are limitations of publication bias, selection bias, inaccuracy of the extracted data, and the problem of misclassification. As defined by [11], these limitations are caused by the visibility problem, where the highly visible works are more pronounced since they are more thoroughly indexed. In this work, the issues rising from this dilemma, publication bias, and selection bias were addressed with the selection of data collection sources. The misclassification and inaccuracy issues were addressed by manually inspecting all of the 601, primary documents, and discarding the documents that were considered out of scope or simply did not contain components that would identify the applied research method as a system-

atic mapping study. On the other side, the misclassification issue of documents not being classified as systematic mapping studies was addressed by using several search sources, and on by relying on keyword/topic word-based searches instead of metadata or simple self-declared classifications, such as the ACM Computing Classification System concepts. During the data mining process, one additional feature was to use prior manually found known works as the measurement points; if the prior known works were captured by the automatic search and collection algorithm, then it was considered to be at least as accurate as the researchers doing manual data collection.

[22] define systematic mapping study to have weaknesses in the lack of synthesis and in-depth analysis especially when compared against other research approaches, broad overall descriptive level, and risk of oversimplifying the results of the studies that were selected as the primary documents. Mapping studies also do not sufficiently provide quality measurement; the quality is highly dependent on the qualities of the researchers. Overall, these limitations cannot be disputed, but they are recognized and there are ways to minimize their risk for the study validity.

As for the other risks in this type of publication, one of the pressing issues is the publication bias (for example [55]). The publication bias refers to the positive, strong or successful outcomes being more likely to be published than negative or inconclusive results, indicating that the publications included in this study are most likely only the reported cases where the systematic mapping studies have been applied successfully, or at least to the degree where significant results have been obtained. This is problematic for the assessment of the usability of the systematic mapping study method in the different domains, but the overall data collection has been conducted from a non-domain-specific search database. The original dataset includes items from several non-peer-reviewed sources such as thesis works or white papers, even though in the actual analysis phase these sources were discarded as unreliable. After filtering out the unreliable publications, the data represents various different scientific domains and nationalities, from international and regional levels. Because of this, it is possible to assume that the data collection method is accurate enough to minimize the publication bias and subsequent visibility bias.

Tools and search strings used in this study also set limitations. Google Scholar updates its database on a daily basis and one cannot replicate searches conducted in this study. Further search strings do not find studies that would today be classified as systematic mapping studies, but have not been named as such when the method itself had not been defined. This is also with the case of systematic literature reviews, which were excluded from this study, although some of them might actually be mapping studies.

VII. CONCLUSION

In this article, a systematic mapping study on the spread and application of systematic mapping studies in the different

TABLE 12. Full list of fields using systematic mapping study methods. Few articles were categorized into two or three fields.

Field	Number of hits
computer science, software engineering	250
health care sciences & services	56
medicine, general & internal	42
robotics	13
education & educational research	9
social issues	9
computer science, hardware & architecture	8
computer science, information systems	7
multidisciplinary sciences	7
business	6
health policy & services	6
computer science, theory & methods	5
medical informatics	5
education, scientific disciplines	4
information science & library science	4
primary health care	4
telecommunications	4
computer science, interdisciplinary applications	3
neurosciences	3
social sciences, interdisciplinary	3
green & sustainable science & technology	2
integrative & complementary medicine	2
management	2
public, environmental & occupational health	2
agricultural economics & policy	1
art	1
computer science, artificial intelligence	1
computer science, interdisciplinary study	1
dentistry, oral surgery & medicine	1
economics	1
engineering, electrical & electronic	1
engineering, manufacturing	1
environmental studies	1
geriatrics & gerontology	1
humanities, multidisciplinary	1
language & linguistics	1
medical laboratory technology	1
pathology	1
pediatrics	1
pharmacology & pharmacy	1
political science	1
psychology	1
psychology, health care sciences & services	1
veterinary sciences	1

scientific domains has been presented based on the data set covering the initial decade after the systematic mapping studies were codified in 2007. The research focus was on the identification of the prominent approaches, protocols, and methods, which are applied in practice in the different scientific domains. Overall, the study identified over six hundred different applications of the systematic mapping study approach, from which 423 peer-reviewed scientific studies were selected for more in-depth analysis.

Overall, the spread and application of some form of systematic mapping study approach was identified from 55 different research domains, with the main domains being computer science and healthcare. With statistical clustering, domains such as information sciences, social studies, and economics were also identified as major application areas.

As based on the observations, the systematic mapping study family is also fragmented, for example, the applied protocol and research procedure correlates strongly with the application domain, and the two major research areas of systematic mapping studies apply different protocol families. Additionally, the practically applied components of the systematic mapping studies indicate that there is a trend of simplification: the more laborious methods such as snowballing, keyword design, or primary candidate grading were identified only in a minority of the conducted studies even though they are considered key aspects of the research method.

As for future work, the development of one protocol to merge all major systematic mapping study trends might be an ideal objective. However, since the frequency of the publication of the systematic mapping studies is increasing constantly, it also implies that the number of different application domains keeps rising. In this context, the next step towards a unified systematic mapping study protocol could be to assess how the major protocol definitions differ from each other, and how applicable they are when applied outside their original domain.

APPENDIX A

COMPLETE LIST OF ACCEPTED ARTICLES

The list of accepted articles can be found from zenodo: <https://zenodo.org/record/6405252> (DOI: 10.5281/zenodo.6405252).

APPENDIX B

COMPLETE LIST OF FIELDS

See Table 12.

REFERENCES

- [1] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf. Softw. Technol.*, vol. 64, pp. 1–18, Aug. 2015.
- [2] D. Badmpudi, C. Wohlin, and K. Petersen, "Experiences from using snowballing and database searches in systematic literature studies," in *Proc. 19th Int. Conf. Eval. Assessment Softw. Eng.*, New York, NY, USA, Apr. 2015, p. 17.
- [3] B. Kitchenham, R. Pretorius, D. Budgen, O. P. Brereton, M. Turner, M. Niazi, and S. Linkman, "Systematic literature reviews in software engineering—A tertiary study," *Inf. Softw. Technol.*, vol. 52, pp. 792–805, Aug. 2010.
- [4] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quart.*, vol. 26, pp. 12–15, Jun. 2002.
- [5] I. D. Cooper, "What is a 'mapping study,'" *J. Med. Library Assoc.*, vol. 104, pp. 76–78, Jan. 2016.
- [6] M. D. J. Peters, C. M. Godfrey, H. Khalil, P. McInerney, D. Parker, and C. B. Soares, "Guidance for conducting systematic scoping reviews," *Int. J. Evidence-Based Healthcare*, vol. 13, no. 3, pp. 141–146, Sep. 2015.
- [7] M. Dijkers, "What is a Scoping Review?" *Knowl. Transl.*, vol. 4, pp. 1–5, Dec. 2015.
- [8] Clarivate Analytics. (2020). *Web of Science Research Areas*. Accessed: Mar. 12, 2021. [Online]. Available: https://images.webofknowledge.com/images/help/WOS/hp_subject_category_terms_tasca.html
- [9] R. Chilton, M. Pearson, and R. Anderson, "Health promotion in schools: A scoping review of systematic reviews," *Health Educ.*, vol. 115, nos. 3–4, pp. 357–376, Jun. 2015.

- [10] S. Mickan, J. K. Tilson, H. Atherton, N. W. Roberts, and C. Heneghan, "Evidence of effectiveness of health care professionals using handheld computers: A scoping review of systematic reviews," *J. Med. Internet Res.*, vol. 15, no. 10, p. e212, Oct. 2013.
- [11] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Keele, U.K., Durham Univ. Joint Rep., Durham, U.K., Tech. Rep. EBSE-2007-01, 2007.
- [12] C. Wohlin, P. Runeson, P. A. da Mota Silveira Neto, E. Engström, I. do Carmo Machado, and E. S. de Almeida, "On the reliability of mapping studies in software engineering," *J. Syst. Softw.*, vol. 86, no. 10, pp. 2594–2610, 2013.
- [13] I. S. Santos, R. M. Andrade, and P. A. Santos Neto, "Templates for textual use cases of software product lines: Results from a systematic mapping study and a controlled experiment," *J. Softw. Eng. Res. Develop.*, vol. 3, no. 1, p. 5, Dec. 2015.
- [14] M. Verdonck, F. Gailly, S. de Cesare, and G. Poels, "Ontology-driven conceptual modeling: A systematic literature mapping and review," *Appl. Ontol.*, vol. 10, nos. 3–4, pp. 197–227, Dec. 2015.
- [15] R. Armstrong, B. J. Hall, J. Doyle, and E. Waters, "'Scoping the scope' of a cochrane review," *J. Public Health*, vol. 33, no. 1, pp. 147–150, Mar. 2011.
- [16] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *Proc. 12th Int. Conf. Eval. Assessment Softw. Eng.*, vol. 8, Jun. 2008, pp. 68–77.
- [17] B. A. Kitchenham, D. Budgen, and O. P. Brereton, "The value of mapping studies: A participant-observer case study," in *Proc. 14th Int. Conf. Eval. Assessment Softw. Eng.*, Swindon, U.K., 2010, pp. 25–33.
- [18] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 151, no. 4, p. 264, Aug. 2009.
- [19] H. Arksey and L. O'Malley, "Scoping studies: Towards a methodological framework," *Int. J. Social Res. Methodol.*, vol. 8, no. 1, pp. 19–32, Feb. 2005.
- [20] D. Rosenstreich and B. Woolscroft, "Measuring the impact of accounting journals using Google scholar and the G-index," *Brit. Accounting Rev.*, vol. 41, no. 4, pp. 227–239, Dec. 2009.
- [21] D. Giustini and M. N. Kamel Boulos, "Google scholar is not enough to be used alone for systematic reviews," *Online J. Public Health Informat.*, vol. 5, no. 2, p. 214, Jun. 2013.
- [22] M. J. Grant and A. Booth, "A typology of reviews: An analysis of 14 review types and associated methodologies," *Health Inf. Libraries J.*, vol. 26, no. 2, pp. 91–108, Jun. 2009.
- [23] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," *J. Syst. Softw.*, vol. 80, no. 4, pp. 571–583, 2007.
- [24] S. Mujtaba, K. Petersen, R. Feldt, and M. Mattsson, "Software product line variability: A systematic mapping study," School Eng., Blekinge Inst. Technol., Karlskrona, Sweden, Tech. Rep., 2008.
- [25] R. Wieringa, N. Maiden, N. Mead, and C. Rolland, "Requirements engineering paper classification and evaluation criteria: A proposal and a discussion," *Requirement Eng.*, vol. 11, no. 1, pp. 102–107, Dec. 2005.
- [26] J. L. Fleiss, B. Levin, and M. C. Paik, *Statistical Methods for Rates and Proportions*. Hoboken, NJ, USA: Wiley, 2013.
- [27] *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [28] J. Fox. (2016). *Polycor: Polychoric and Polyserial Correlations*. Accessed: Jan. 30, 2021. [Online]. Available: <https://CRAN.R-project.org/package=polycor>
- [29] F. Drasgow, "Polychoric and polyserial correlations," *Encyclopedia Stat. Sci.*, vol. 7, pp. 68–74, 1986.
- [30] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandin. J. Statist.*, vol. 4, pp. 65–70, Dec. 1979.
- [31] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
- [32] A. Knutas, A. Hajikhani, J. Salminen, J. Ikonen, and J. Porras, "Cloud-based bibliometric analysis service for systematic mapping studies," in *Proc. 16th Int. Conf. Comput. Syst. Technol.*, 2015, pp. 184–191.
- [33] K. Hornik and B. Grün, "TopicModels: An R package for fitting topic models," *J. Stat. Softw.*, vol. 40, no. 13, pp. 1–30, 2011.
- [34] C. Sievert and K. Shirley, "LDAvis: A method for visualizing and interpreting topics," in *Proc. Workshop Interact. Lang. Learn., Visualizat., Interface*, 2014, pp. 63–70.
- [35] D. Blei, L. Carin, and D. Dunson, "Probabilistic topic models," *Commun. ACM*, vol. 55, no. 4, pp. 77–84, Nov. 2010.
- [36] D. Mimno, H. M. Wallach, E. Talley, M. Leenders, and A. McCallum, "Optimizing semantic coherence in topic models," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2011, pp. 262–272.
- [37] M. E. Roberts, B. M. Stewart, and D. Tingley. (2017). *STM: R Package for Structural Topic Models*. Accessed: Jan. 30, 2021. [Online]. Available: <http://www.structuraltopicmodel.com>
- [38] D. D'Amato, N. Droste, B. Allen, M. Kettunen, K. Lähinen, J. Korhonen, P. Leskinen, B. D. Matthies, and A. Toppinen, "Green, circular, bio economy: A comparative analysis of sustainability avenues," *J. Cleaner Prod.*, vol. 168, pp. 716–734, Dec. 2017.
- [39] A. Knutas, *Increasing Beneficial Interactions in a Computer-Supported Collaborative Environment*. vol. 718. Lappeenranta, Finland: Lappeenranta University of Technology, 2016.
- [40] B. Penzenstadler, A. Raturi, D. Richardson, C. Calero, H. Femmer, and X. Franch, "Systematic mapping study on software engineering for sustainability (SE4S)," in *Proc. 18th Int. Conf. Eval. Assessment Softw. Eng.*, 2014, pp. 1–14.
- [41] C. Wang and D. M. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, p. 448.
- [42] SCImago. (2017). *Scimago Journal & Country Rank*. Accessed: Aug. 7, 2020. [Online]. Available: <http://www.scimagojr.com/>
- [43] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proc. 18th Int. Conf. Eval. Assessment Softw. Eng.*, New York, NY, USA, 2014, p. 38.
- [44] K. Petersen and C. Gencel, "Worldviews, research methods, and their relationship to validity in empirical software engineering research," in *Proc. Joint Conf. 23rd Int. Workshop Softw. Meas.*, Washington, DC, USA, Oct. 2013, pp. 81–89.
- [45] S. Baumgart and J. Froberg, "Functional safety in product lines—A systematic mapping study," in *Proc. 42th Euromicro Conf. Softw. Eng. Adv. Appl. (SEAA)*, Limassol, Cyprus, Aug. 2016, pp. 313–322.
- [46] A. Gabriel, D. Monticcolo, M. Camargo, and M. Bourgault, "Creativity support systems: A systematic mapping study," *Thinking Skills Creativity*, vol. 21, pp. 109–122, Sep. 2016.
- [47] C. Moraes, L. Philippsen, H. Lirani, L. Yamanaka, D. Rosim, and E. E. Filho, "Systematic mapping study in small business: The quest for contemporary understanding," *Proc. Social Behav. Sci.*, vol. 143, pp. 916–920, Aug. 2014.
- [48] E. Souza, D. Castro, D. Vitória, I. Teles, A. L. I. Oliveira, and C. Gusmão, "Characterizing user-generated text content mining: A systematic mapping study of the Portuguese language," in *Proc. 12th Int. Conf.*, vol. 444. Tomar, Portugal: Springer, 2016, pp. 1015–1024.
- [49] B. Kitchenham, "Procedures for performing systematic reviews," Keele Univ., Keele, U.K., Durham Univ. Joint Rep., Durham, U.K., Tech. Rep. TR/SE-0401, 2004.
- [50] B. A. Kitchenham, D. Budgen, and O. Pearl Brereton, "Using mapping studies as the basis for further research—A participant-observer case study," *Inf. Softw. Technol.*, vol. 53, pp. 638–651, Jun. 2011.
- [51] D. Levac, H. Colquhoun, and K. K. O'Brien, "Scoping studies: Advancing the methodology," *Implement. Sci.*, vol. 5, p. 69, Sep. 2010.
- [52] C. P. Dancy and J. Reidy, *Statistics Without Maths for Psychology*. London, U.K.: Pearson, 2007.
- [53] M. Petticrew, "Why certain systematic reviews reach uncertain conclusions," *BMJ Clin. Res.*, vol. 326, pp. 756–758, Apr. 2003.
- [54] J. C. van Niekerk and J. D. Roode, "Glaserian and straussian grounded theory: Similar or completely different?" in *Proc. Annu. Res. Conf. South Afr. Inst. Comput. Sci. Inf. Technol.*, New York, NY, USA, 2009, pp. 96–103.
- [55] S. Duval and R. Tweedie, "Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis," *Biometrics*, vol. 56, pp. 455–463, Jun. 2000.



ERNO VANHALA received the Doctor of Science degree in software engineering from the Lappeenranta University of Technology (LUT), Lappeenranta, Finland, in 2015. He was working with Tampere University as a Web Designer and Developer, but returned to academia, in 2020. He is working as a University Lecturer with LUT. He has published international research articles on topics, such as business, development innovation, and engineering aspects of computer games. His current research interests include computer game start-ups and their business models. Besides business issues, he is also mesmerized by the open source phenomenon and web-based software. He is a member of the Finnish Centre for Open Systems and Solutions (COSS); the Vice-Chairperson of Uskon-tojen uhrien tuki UUT ry; and a merited Teacher, having received several awards on teaching different software engineering topics.



ANTTI KNUTAS received the Doctor of Science degree in technology from the Lappeenranta University of Technology, in 2016, with a focus on communication software. He is an Assistant Professor. He is currently working as a Postdoctoral Researcher with the Lappeenranta University of Technology. He is also contributing to the nails project, an open source effort to produce. His main research interests include computer-supported cooperative work, collaboration, gamification, and social networks analysis. He has received several awards from his work on social data analytics and educational collaboration research.



JUSSI KASURINEN received the Doctor of Science degree in technology. He is an Associate Professor with LUT University, specializing in software engineering and software testing, and an Adjunct Professor of entertainment software engineering. His current research work with LUT University focuses on software processes, software quality assurance, games as software, digitalization, and digital economy. He has also been working with software testing, test processes, software quality, and computer science education. He has been doing research collaboration with over 40 different software developing companies in Finland and Northern Europe; and has also published books on different topics, such as testing and quality assurance, and esoteric programming languages in Finnish.



ANTTI HERALA received the Ph.D. degree in software engineering from the Lappeenranta University of Technology, Finland, in 2018. The author has published work on education in conferences, concentrating on the flipped classroom method and its benefits to students and educators alike. He is currently employed by MP Soft, Finland, where he is the Chief Technology Officer.

...