

Review

The Impact of Controlled Vocabularies on Requirements Engineering Activities: A Systematic Mapping Study

Arshad Ahmad ^{1,2}, José Luis Barros Justo ³ , Chong Feng ^{1,*} and Arif Ali Khan ⁴

¹ School of Computer Science & Technology, Beijing Institute of Technology, Beijing 100081, China; yaarshad@uoswabi.edu.pk or arshad.ahmad@fecid.paf-iast.edu.pk

² Department of IT and Computer Science, Pak-Austria Fachhochschule: Institute of Applied Sciences and Technology, Mang Khanpur Road, Haripur 22620, Pakistan

³ School of Informatics (ESEI), University of Vigo, 32004 Ourense, Spain; jbarros@uvigo.es

⁴ College of Computer Science & Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China; arif.khan@nuaa.edu.cn

* Correspondence: fengchong@bit.edu.cn

Received: 24 August 2020; Accepted: 27 October 2020; Published: 2 November 2020



Abstract: Context: The use of controlled vocabularies (CVs) aims to increase the quality of the specifications of the software requirements, by producing well-written documentation to reduce both ambiguities and complexity. Many studies suggest that defects introduced at the requirements engineering (RE) phase have a negative impact, significantly higher than defects in the later stages of the software development lifecycle. However, the knowledge we have about the impact of using CVs, in specific RE activities, is very scarce. Objective: To identify and classify the type of CVs, and the impact they have on the requirements engineering phase of software development. Method: A systematic mapping study, collecting empirical evidence that is published up to July 2019. Results: This work identified 2348 papers published pertinent to CVs and RE, but only 90 primary published papers were chosen as relevant. The process of data extraction revealed that 79 studies reported the use of ontologies, whereas the remaining 11 were focused on taxonomies. The activities of RE with greater empirical support were those of specification (29 studies) and elicitation (28 studies). Seventeen different impacts of the CVs on the RE activities were classified and ranked, being the two most cited: guidance and understanding (38%), and automation and tool support (22%). Conclusions: The evolution of the last 10 years in the number of published papers shows that interest in the use of CVs remains high. The research community has a broad representation, distributed across the five continents. Most of the research focuses on the application of ontologies and taxonomies, whereas the use of thesauri and folksonomies is less reported. The evidence demonstrates the usefulness of the CVs in all RE activities, especially during elicitation and specification, helping developers understand, facilitating the automation process and identifying defects, conflicts and ambiguities in the requirements. Collaboration in research between academic and industrial contexts is low and should be promoted.

Keywords: controlled vocabularies; requirements engineering; software development; evidence-based software engineering; systematic mapping study

1. Introduction

Requirements engineering (RE) is the foremost, human-centric, and crucial phase of the software development lifecycle, concerned with adequately eliciting, analyzing, validating, and managing

user requirements. In real-world scenarios, it is quite hard to correctly fulfil all the requirements, which ultimately leads to poor quality software or project failure [1].

A controlled vocabulary is an organized collection of units of significance, i.e., terms that have determined and well-known meaning, without duplicates (synonyms) that can cause ambiguities or misunderstandings. Their purpose is to organize information (data pieces) in a structured manner, provide consistency, indicate semantic relations, making it easy to classify, query, and retrieve data [2–4].

Although developers have extensive knowledge of software development methods, they often ignore the domain of the problem [5]. Lacking such type of knowledge may lead to problems such as missing important information or specifying ambiguous, contradictory, or incomplete requirements [6]. To reduce or mitigate such problems, it is essential to use adequate methods, techniques, and tools that can effectively deal with domain concepts and their relationships, helping the developers to understand their intricacies. Controlled vocabularies (CVs) may be used as one of these effective strategies to reduce problems related to domain concepts in RE, since they reduce ambiguities and complexities.

There are some previous studies that have investigated the use of CV in RE [7–11]. However, none of these secondary studies captures all aspects, impacts and evidence that interest us. Our goal is to explicitly identify and categorize the different types of CV available, and the support they provide to activities in the RE phase of software development.

For this work, we have chosen the systematic mapping study (SMS) as a research method [12–14], to identify and classify the available empirical evidence about the use of CVs in RE. The SMS research approach has been used successfully on different topics within the scope of RE [15–21].

The research method will help to get a deeper knowledge about CVs and their impact in RE activities, helping to:

- Detect research gaps (future research opportunities);
- Aid decision-making (practitioners) when selecting a CV or a tool;
- Better plan the RE phase, avoiding pitfalls.

The specific contributions of this work are to:

- C1: Identify and classify the CVs (RQ1) used in activities related to the RE phase (RQ2) of software development;
- C2: Identify the impact of CVs on the development process and the final product (RQ3);
- C3: Identify some demographic data such as active researchers, organizations and countries, and the most frequent publication venues (DQs);
- C4: Gather dispersed evidence providing a centralized source to facilitate research.

The remainder of this paper is organized as follows: Section 2 presents the background and related work whilst Section 3 details the research method. Results and discussion are presented in Sections 4 and 5 deals with the validity threats. Finally, Section 6 presents the conclusions and proposals for future work.

2. Background and Related Work

Requirements engineering is the earliest process in software development, and therefore, the most crucial, since the following phases depend on it. An unnoticed defect during this phase has important consequences in later development phases [1]. On the other hand, the activities carried out during the RE have a strong component of human labor and are prone to failures and defects. It is necessary to have support methods, techniques and tools that guarantee the quality of the outcomes of the RE, reducing as much as possible the presence of inconsistencies, ambiguities, omissions or defects in the final specifications of the requirements [6].

In the last two decades, there has been a surge in the use of CVs in RE for improving the overall quality of both the development process and the final software product (system) [7–9,22–25]. For instance, the notion of ontologies has been used in RE, to improve the completeness and correctness

of requirements specification, to assist in comprehending requirements, to help in modelling domain knowledge and to effectively manage requirements change [22].

2.1. Controlled Vocabulary

A controlled vocabulary is an organized collection of units of significance, i.e., terms that have a determined and well-known meaning, without duplicates (synonyms) that can cause ambiguities or misunderstandings. Their purpose is to organize information (data pieces) in a structured manner, provide consistency, indicate semantic relations, making it easy to classify, query, and retrieve data [2–4]. The most frequent examples of CVs can be categorized as ontologies, taxonomies, thesauri and, lately, folksonomies. Other examples of CVs include header–subject lists, classification schemes or specialized glossaries and dictionaries [26]. However, all these CV variations can be considered as included in the previous categories. Support tools to CV's applications make extensive use of NLP (natural language processing) and knowledge management techniques [23,27].

Ontologies are formal representations of concepts, ideas or objects, pertaining to a specific domain, as well as all the relationships between those concepts [28]. The ultimate goal of ontologies is to facilitate automated reasoning and, by doing that, provide a deeper understanding of the domain.

A taxonomy is a classification of things or concepts, often with an explicitly imposed hierarchy and a set of principles that underlie such classification [28]. Search engines exploit taxonomies to increase the precision or relevance of query results, as well as to speed up the search itself.

A thesaurus groups together words with similar (synonyms) and opposite (antonyms) meanings [29]. Thesauri are reference tools that provide users with a choice of the most adequate terms (words) to associate with an idea or concept. The combination of these “most adequate terms” leads to precise, fit and apt definitions of an idea or concept and avoids misunderstandings.

A folksonomy is a system of classification of terms or concepts created from a collaborative labelling work. Folksonomy is also known as social tagging, collaborative tagging, social rating and social bookmarking. The fundamental difference with respect to a taxonomy is that in the latter the terms and their meanings are pre-established, whereas in folksonomies they constantly evolve and change in the classification scheme based on the contributions of user groups. The most frequent applications of folksonomies are in the areas of social networks and websites, to locate and recover multimedia content [24].

Controlled vocabularies make extensive use of the above concepts and are widely used in the requirements engineering domain [18,25].

2.2. RE Activities

RE is one of the most critical and complex activities in the software development process because decisions about what should be developed, when and how depend on it [30]. RE deals with Elicitation, Analysis and Negotiation, Specification, Validation and Change Management [31].

The Elicitation phase of requirements implies understanding the general domain of the problem and the specific problem to be solved, identifying the main actors and stakeholders, knowing the wishes and limitations of the organization, and gathering all the information relevant to starting the project.

The analysis and negotiation are related to the high-level requirements from the previous elicitation and specification tasks. The objective of this activity is to produce a set of complete and consistent requirements.

The requirements specification produces documents with complete, consistent and non-ambiguous descriptions of the actors' needs. These descriptions must fully identify the interactions of the actors with the system to be developed, the necessary inputs and the desired outputs.

Finally, the aim of the validation phase is to verify that requirements have been described entirely and acceptably. The inputs to this process include the previous document with the requirements specifications, inner rules and protocols from the organization, legal issues and implicit knowledge of the developers.

The management process, as suggested by Kotonya and Sommerville [6], involves all the above tasks to deal with new requirement's requests or changes to existing ones.

2.3. Related Works

Parreira et al. [7] conducted a Systematic Mapping Study (SMS) in 2015 focused on domain ontologies in the context of RE. The goal of their work was to identify and get a detailed insight into the usage of domain ontologies in RE activities. Their search strategy included only an automated search in electronic databases and was run on June 2014. They selected and analyzed 67 primary studies to answer four research questions. Their key findings were: (1) in the context of RE, ontologies are related to requirements elicitation, analysis and verification, conflict identification and analysis, and unification among requirements formalisms; (2) the majority of the studies (59%) were found to be related to requirements elicitation and analysis; (3) there is a lack of evaluation studies on the usage of ontologies in RE; (4) most of the selected studies (80.6%) reported about general purpose software applications.

In 2016, Demerval et al. [8] conducted an SLR focused on the applications of ontologies in RE. The main goal of their work was to get a more profound and better understanding of the support that ontologies bring to RE phases. The authors used the classification of RE process proposed by Kotonya and Sommerville [6]. Automated search in databases was the only search strategy used and covered the period from January 2007 to October 2013. In their work, they analyzed the requirements modelling styles, the type of supported requirements and the ontology languages that have been used. Data were extracted from 67 selected primary studies to answer a set of seven research questions. The main findings of their study were: (1) the most frequently reported benefits from using ontologies in RE were reducing ambiguity, incompleteness and inconsistency of requirements (57% of the studies), followed by requirements management and evolution (36%) and domain knowledge representation (27%); (2) ontologies support a great diversity of RE modelling styles, mainly textual requirements and UML; (3) studies addressed only functional requirements (52%) or a combination of functional and non-functional (45%), only 3% of the studies reported just non-functional requirements; (4) although a great variety of ontologies were found, none of them have been widely adopted by the research community. These findings are limited by the context of selected papers (78% academic and 22% industrial setting). The work from Dermeval et al. shows a very high level of quality and a sound research method.

Anu et al. [9] conducted an SLR study to identify, analyze and classify possible reasons of human errors that occur during requirements engineering and their impact on software quality. They only used an automated search strategy in electronic databases, conducted from 2006 to October 2014. The data were extracted from 38 selected primary studies. The authors developed a Human Error Taxonomy (HET), based on their results and by extending a previous Requirement Error Taxonomy [32]. The HET is based on Reason's taxonomy that comprised of slips, lapses, and mistakes. The mapping of the HET to all phases of RE revealed that there is a lack of evidence about errors for requirements verification and management phases. Their main contributions were: (1) introducing and applying the human error research to a novel software quality improvement domain via interaction between RE experts and cognitive psychology researchers; (2) identifying and classifying the human errors, facilitating a systematic way to comprehend and avoid those that arise during human-centric activities of requirements engineering.

An SLR focused on the Requirements Change Management (RCM) process was conducted by Jayatilke et al. [10] in 2018. Their study aimed to investigate the causes of requirements changes, processes and techniques used for RCM and how the organizations deal with decision making during RCM process. They ran an automated search in six electronic databases but did not report the search date or the covered period. Data were extracted from 184 primary studies. Their main findings were: (1) The causes of RCs can be divided into trigger events and uncertainties, and in five different areas. This classification can lead to better planning that will ensure a better success rate for the project;

(2) The key processes of RCM are: change identification, change analysis and change effort estimation; (3) Taxonomies and classifications are frequently used to change identification; (4) Decision making can be contradictory at different levels of the organization. This can cause a contradictory understanding of the change between the business and IT counterparts. Their general conclusion was that RCM is an elusive target to achieve and that there are many ways to tackle it.

Pachecho et al. [11] conducted an SLR in 2018 focused on analyzing the maturity of techniques used for requirements elicitation. The goal of their work was to collect and assess the available evidence of requirements elicitation techniques to aid the software analysts in picking the suitable technique for eliciting software requirements. The search strategy consisted of an automated search only, ran in seven electronic databases and covering a period from 1993 to 2015. The data for the SLR were extracted from 140 selected primary studies to answer two research questions. They concluded that the most widely used elicitation techniques identified were: traditional techniques, e.g., interviews, task analysis and questionnaires (27%). Similarly, modelling techniques, e.g., goal-based, scenarios, and business process models (21%) and collaborative techniques, including workshops focus groups and brainstorming (9%). Moreover, cognitive techniques, e.g., ontology, card sorting and repertory grid (7%), and agile techniques, e.g., user stories, mind mapping and group storytelling (7%) among others. Nevertheless, the selection and suitability of these techniques depend upon the stakeholders' characteristics, resources available, context and the problem at hand to be solved. The findings from the selected studies came from a well-balanced context (55.8% academic, 38.8% industrial, and 5.4% mixed settings). The authors followed a sound research method, and, because of this, their work shows a high level of quality.

Unlike the works cited above, our study was not limited to any pre-established CV or specific RE activity. On the other hand, to ensure wider coverage of potential sources of evidence, we included two complementary search strategies: automatic search in electronic data sources (EDS) and a backward (based on references) and forward (based on citations) snowballing [33]. The period considered covered until June 2019 (searches were conducted in July 2019). Regarding the number of primary papers included, our study considered a similar amount to the average support level provided by the related works (99 primary works), but 36% more than the average (57 primary papers) of those works considering all the RE activities. A summary of these facts is shown in Table 1.

Table 1. Summary of differences between related work and this study.

Related Work/Year	Coverage	Type of CV	RE Activities	Support Level
Parreira/2015	June 2014	Ontologies	All	67
Demerval/2016	2007–October 2013	Ontologies	All	67
Anu et al./2018	2006–October 2014	Taxonomies	All	38
Jayatilleke/2018	Not reported	Taxonomies	Change Management	184
Pachecho/2018	1993–2015	Ontologies	Elicitation	140
This study/2019	June 2019	All	All	90

3. Research Method

We conducted an SMS, following the guidelines in [34], to identify and classify the use of controlled vocabularies in requirements engineering activities during the software development process. The steps of the SMS research approach/method included the following main activities:

- Planning the SMS research method (the protocol):
 - Definition of the goal and the set of research questions;
 - Specification of the strategies for: search, selection and data extraction processes;
 - Consideration of any possible validity threats;
 - Tasks assignment (roles and responsibilities of every researcher).
- Conducting the SMS method (executing the protocol):

- Searching for primary papers;
- Selection of the relevant primary papers;
- Data extraction and thorough analysis of the selected primary papers to produce a classification schema (the map).
- Reporting the results

A previous protocol was published in the Computing Research Repository (CoRR) in 2017 (<https://arxiv.org/abs/1704.00822>), and since then the goal and research questions have changed a bit; that is why a new, slightly different, protocol was designed for this study. The following subsections offer detailed information regarding the key activities of this new protocol.

3.1. Goal and Research Questions

The main research goal of this SMS is:

To identify and categorize/classify the type of controlled vocabularies used during the requirements engineering phase and to identify their impact on software development.

This research goal is broad enough to allow for subdividing it into the set of research questions (RQs) detailed in Table 2.

Table 2. Description of the research questions.

Research Question	Description
RQ1: Which types of CVs are reported?	What the CVs are based on? (e.g., Folksonomies, Ontologies, Taxonomies, Thesauri or Other)
RQ2: In which RE activities have the CVs been used?	The RE activity: Elicitation, Specification, Analysis, Validation, Change Management or Other.
RQ3: Which aspects of the software development process, or of the final product, were impacted by the use of the CVs?	In the first phase the name of the aspect will be extracted, verbatim, as it appears in the original primary document (for example, productivity, quality, development time, ease of maintenance, error reduction, automation, etc.). Due to the possible variety of terms used, a thematic analysis will be carried out in a second phase to group all these terms into a set of representative categories.

We also gathered information to answer some interesting demographic questions (DQs) related to the identification of the most active researchers, organizations and countries, as well as the top publication venues. Table 3 presents a brief description of every DQ.

Table 3. Description of the demographic questions.

Demographic Question	Description
DQ1: Who are the most active researchers?	All authors, ordered by the number of papers.
DQ2: Which are the most active organizations?	Based on the affiliations of all authors.
DQ3: Which are the most active countries?	Based on the affiliations of all authors.
DQ4: Which are the top publication venues?	Type (Conference, Journal or Workshop) and the Name of the publishing venue.

3.2. Search Strategies

The EBSE (evidence-based software engineering) paradigm relies on the process of gathering all the available published empirical evidence. As it was impossible to ensure that all existing empirical evidence was found, we had to reduce the validity threat of not considering relevant sources [35], for this reason, two search strategies were included, to complement each other, retrieving the highest

number (recall) of potentially relevant studies (precision). Experts on the CVs and RE areas validated the search strategies.

The automated search strategy was conducted in four different electronic data sources (namely; ACM Digital Library (DL), IEEE Xplore, SCOPUS and Web of Science) [36,37]. These electronic data sources (EDS) cover the principal venues (Conferences, Workshops and Journals) in the software engineering field [38,39]. The search string was built from the key terms derived from our defined set of different RQs, keywords from papers retrieved via pilot searching and from a list of possible synonyms.

The original key terms were combined using the AND logical operator, whereas their synonyms used the OR operator. Several search rounds were performed until the authors reached the best balance between precision and recall measures. Table 4 depicts the final search strings, tailored to each of the EDS. This automated search strategy was run on 9 July 2019, retrieving 901 results.

Table 4. Search strings tailored to each electronic data source (EDS).

Source	Search String
ACM DL	(acmdlTitle:(+control* +vocabulary*) OR acmdlTitle:(folksonom* thesaur* ontolog* taxonom*)) AND ((acmdlTitle:(+requirement +engineering) OR recordAbstract:(+requirement +engineering) OR keywords.author.keyword:(+requirement +engineering))
IEEE Xplore	((("Document Title": "controlled vocabulary") OR ("Document Title": "folksonom*" OR "Document Title": "ontolog*" OR "Document Title": "taxonom*" OR "Document Title": "thesaur*")) AND ((("Document Title": "requirements" AND "Document Title": "engineering") OR ("Abstract": "requirements" AND "Abstract": "engineering") OR ("Author Keywords": "requirements" AND "Author Keywords": "engineering"))
SCOPUS	TITLE ((control* AND vocabulary*) OR folksonom* OR ontolog* OR taxonom* OR thesaur*) AND TITLE-ABS-KEY (requirement* AND engineering) AND SUBJAREA (comp) AND (SRCTYPE (j) OR SRCTYPE (p)) AND (LANGUAGE (english))
Web of Science	TI = ((control* AND vocabulary*) OR folksonom* OR ontolog* OR taxonom* OR thesaur*) AND TS = (requirement* AND engineering) AND LANGUAGE: (English) Refined by: WEB OF SCIENCE CATEGORIES: (COMPUTER SCIENCE SOFTWARE ENGINEERING)

In addition to the automated search strategy, backward and forward snowballing [40,41] were performed as a complementary search strategies. All the 70 selected studies from the automated search strategy were used as initial seeds for the snowballing process. In every iteration of the snowballing it is crucial to include only the relevant studies; to that end we followed a top-down sequential process to select only the related studies for every new iteration. To double-check that relevance, the authors also conducted the data extraction to be sure that the pre-selected primary studies were able to answer the RQs. Only those papers that successfully passed this data extraction were used as seeds for the next iteration of snowballing.

The snowballing process ended at the fourth iteration, after retrieving 1446 papers. A set of 20 primary studies were selected. To reduce bias, two authors independently performed the selection of studies, and a third author reviewed the data of every iteration, combined the individual results and checked for disagreements or ambiguities. Tables 5 and 6 show the outcomes of these two search strategies.

Table 5. Automated search results.

Source	Works Retrieved
ACM DL	75
IEEE Xplore	126
SCOPUS	580
Web of Science	120
Total=	901

Table 6. Snowballing search results.

SB Iteration	Seeds	Citations
First	70	1176
Second	14	162
Third	5	99
Fourth	1	9
Total=	90	1446

3.3. Paper Selection: Inclusion/Exclusion Criteria

The studies selection process included mainly two tasks: proper definition of the include and exclude criteria and the application of these defined criteria to choose the most relevant primary published papers [42,43].

As we know that the inclusion and exclusion criteria are totally two opposite actions, the authors decided to concentrate the efforts on exclusion, by properly defining a set of criteria that is both objective and subjective. The former does not cause any type of threat to the validity, and hence, its application is quite simple. Applying these first exclusion criteria, particularly those concerning language and duplicity, allowed the authors to eliminate irrelevant content quickly. The following list contains the objective exclusion criteria applied to the retrieved published research studies:

- EC1: Papers written in other languages except English;
- EC2: Short published research studies (less than four pages in length);
- EC3: Research studies not published in some peer-reviewed venues;
- EC4: Not a primary research study (secondary and tertiary studies, if any, were considered in the Related Work section of this study);
- EC5: Grey literature (books, slide presentations, forewords, talks, etc.);
- EC6: PhD or Master Theses, under the assumption that relevant publications resulting from them were already published as research papers on peer-reviewed venues;
- EC7: Duplicate reports of the same study (consider only the most recent one),

The subjective criteria are the trickiest to handle. They possibly can bring bias into the research study, and a predefined protocol must be used to lessen this threat as much as is viable. Contrastingly, the application of these criteria produces the greatest decrease in the number of research studies to deem as relevant. The authors applied mainly two exclusion criteria stated below:

- Not Focus: Not relevant to the application of controlled vocabularies in software development;
- Out of Scope: Not relevant to any of the requirements engineering phase of software development lifecycle,

Any research study not excluded by the aforementioned criteria was finally comprised in the set of *selected primary research papers*.

A top-down approach was followed to apply these criteria. At first, some particular metadata information such as the title, abstract and keywords was taken into consideration. If this information were not sufficient to exclude the research study at hand, then the authors comprehensively reviewed the full text, specifically the Introduction (problems and contributions of the study), Results and Conclusions sections.

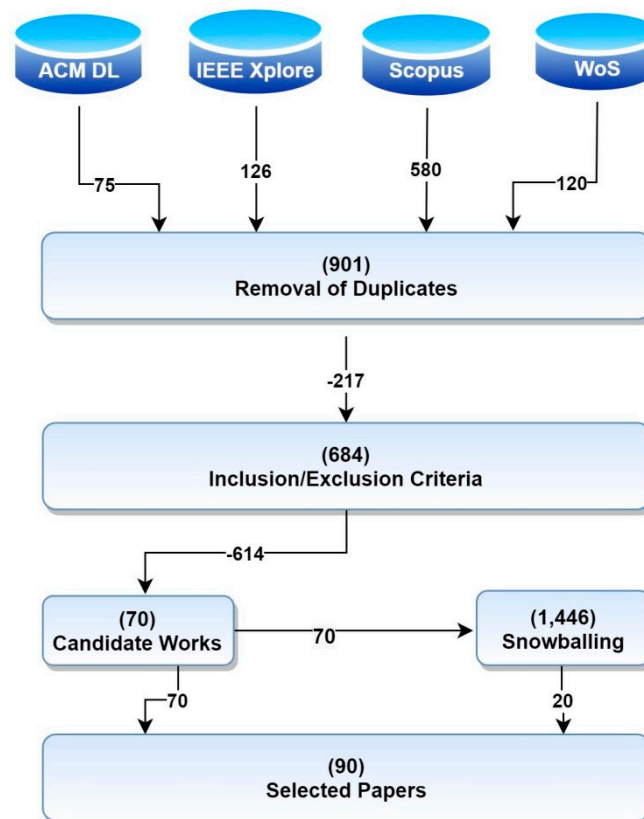
To perform the papers selection process, the papers were equally distributed among authors. A review meeting was held to ensure the smoothness of the process, no relevant paper is overlooked and no irrelevant paper is included.

To deal with disagreements, we applied the inclusive criteria proposed in [34] and detailed Table 7. We excluded a paper only when both reviewers agreed (category “F”) or considered the paper as borderline (category “E”).

Table 7. Dealing with Disagreements.

		Reviewer X		
		Include	Uncertain	Exclude
Reviewer Y	Include	A	B	D
	Uncertain	B	C	E
	Exclude	D	E	F

Figure 1 reflects the flows of the search and selection processes applied, and show the outcome of every task. A final set of 90 primary research papers was selected for this research study (a list with full bibliographic references is provided in Appendix A).

**Figure 1.** Summary of the Search and Selection processes.

3.4. Data Extraction

To lessen the bias in the whole extraction process, a Data Extraction Form was designed by the second author in spread sheet format (available online at [44]). The other authors thoroughly reviewed and agreed upon the Data Extraction Form (DEF) before the formal data extraction process started. The usage of the DEF provides a clear and consistent method to perform the extraction process of the SMS study [45,46].

To accomplish the extraction process task, we divided the set of selected primary papers into two halves as described below:

1. First half: This first half was assigned to reviewers R1 (first author) and R2 (second author) via blind assignment. The reviewers assessed the work independently, and after completion, resolved any differences to produce an agreed dataset.
2. Second half: This second half was assigned to reviewers R1 and R3 (third author), as a blind assignment.

Each pair of the designated reviewers filled out the DEF independently. When any sort of discrepancies arose, an agreement meeting was held, with the presence of a third reviewer, until all sorts of disagreements were resolved [47].

We inspected the Title, Abstract and Conclusions from every primary selected paper to extract the relevant data to answer our set of RQs. If we did not find the needed data in these sections, we inspected the full text of the paper.

4. Results and Discussion

In this section, we present and briefly discuss the outcomes of the SMS study. For each of the RQs we begin with a summary of the most notable outcomes, a brief discussion regarding the most pertinent aspects and, based on them, a suggestion of some explanatory hypotheses. Due to limitations in the number of pages of this study, the detailed results are offered as complementary material, available online at [44].

4.1. RQ1: Which Types of CVs Are Reported?

To accomplish the data extraction process, we initialized with a pre-established classification scheme to group the types of reported CVs. The classification scheme includes Folksonomies, Ontologies, Taxonomies and Thesauri. We only found direct evidence for Ontologies and Taxonomies. Ten selected papers (S12, S16, S29, S31, S33, S52, S73, S74, S86 and S88) reported the use of Thesauri; but they focused on Ontologies, considering the thesaurus only as part of the reported ontology. No direct evidence was found for folksonomies or other terms that can be considered as CVs.

We identified 79 primary papers related to the use of ontologies in RE and eleven related to taxonomies, as depicted in Table 8. Appendix A shows the full list of selected papers.

Table 8. Papers reporting different types of CVs.

Primary Paper	Type of CV	Support Level
S1, S2, S3, S4, S5, S8, S9, S10, S11, S12, S15, S16, S17, S18, S19, S22, S23, S24, S25, S27, S28, S29, S30, S31, S32, S33, S34, S35, S36, S37, S38, S40, S41, S42, S43, S44, S45, S46, S47, S48, S51, S52, S53, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S68, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S84, S85, S86, S87, S88, S89, S90.	Ontology	79
S6, S7, S13, S14, S20, S21, S26, S39, S49, S50, S83.	Taxonomy	11

Most of the mentions to ontologies in RE were related to domain ontologies for facilitating the specification of requirements, elicitation and requirements reuse. These results are not surprising as the aim of ontologies is to reduce ambiguities and gain deeper insights of a certain domain. Some examples of the frequent uses of domain ontologies reported in the selected studies were: security requirements (S18, S38, S60, S62, S66, S90), reducing ambiguities (S10, S18, S36, S41, S58), improving understanding (S08, S23, S27, S32, S45, S50, S65), detecting and resolving conflicts (S02, S19, S43, S44, S49, S72) and specifying requirements (S04, S78, S79, S89) among other uses.

The main categories of impacts of using taxonomies in RE were related to detecting faults and defects (S07, S26), scenarios (S14) and privacy goals (S06) for eliciting and specifying software requirements.

It is worth mentioning that our search strategies did find papers about folksonomies, but they were focused on web applications, tagging, indexing, sentiment analysis and social media. They were not explicitly related to any RE activity, so we discarded those papers because of the exclusion criterion “Out of scope”.

4.2. RQ2: In Which RE Activities Have the CVs Been Reported?

In order to provide meaningful answers to this RQ, we mapped the RE activities reported in the primary studies to those proposed in [6]: Elicitation, Specification, Analysis, Validation,

Change Management and others. We have found that most of the mentions refer to the specification (26.6%), and elicitation activities (25.7%). CVs have also been employed in the Analysis, Validation, Change Management and others, but to a considerably lesser degree. Figure 2 shows the quantity of mentions to every activity in the selected set of papers.

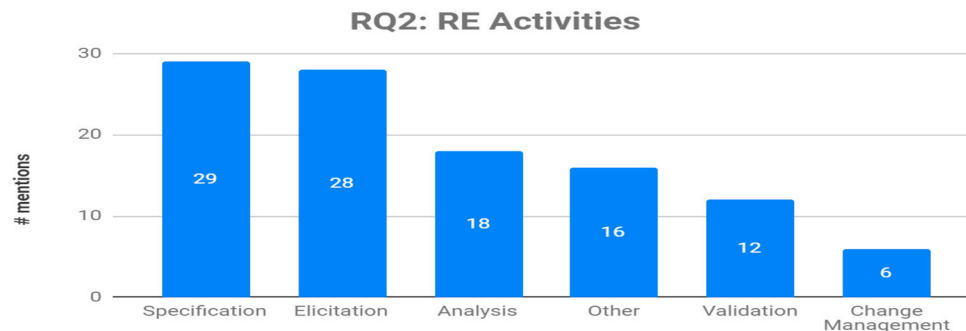


Figure 2. Mentions of requirements engineering (RE) activities where CVs have been used.

It is not surprising that the requirement specification is the activity in which CVs are used most frequently. During this activity, the two types of CVs identified in RQ1 were used. The CVs are used mainly to write the earlier versions of the requirements, informally, during the initial phase of Elicitation, and more formally, during the final step of producing a formal Specification. All papers highlight the need for clear, objective and unambiguous writing. Thus, the papers that report the use of the CVs in the Analysis, Change Management, Validation or Other activities also refer, although indirectly, to the drafting of the requirements specifications.

The papers selected for this study focused on elicitation of requirements from the identified stakeholders or software requirements documents, which in turn can include a variety of sources: interviews, questionnaires, workshops with brainstorming, technical documentation review, simulations, modelling and organizational analysis techniques (e.g., Strengths–Weaknesses–Opportunities–Threats analysis).

The majority of the selected papers related to requirements analysis focused on requirements conflict identification and resolution tasks.

The selected papers reported focused on quality assurance activities of testing and finding faults in requirements. The selected papers related to requirements change management utilized ontologies and taxonomies for handling requirements volatility. Finally, the other activities reported in the selected studies refer to requirements traceability, requirements recovery, requirements engineering methodology (process, methods, and tools) and configuration management among others.

To facilitate further analysis to interested researchers or practitioners, Table 9 detailed the selected papers reporting each RE activity.

Table 9. Papers reporting RE activities.

Primary Papers	RE Activity	Support Level
S4, S5, S6, S14, S17, S18, S20, S22, S23, S25, S26, S30, S33, S34, S36, S40, S41, S45, S48, S49, S58, S63, S64, S65, S75, S76, S78, S85, S90.	Specification	29
S2, S8, S12, S13, S16, S27, S28, S29, S32, S34, S39, S40, S55, S59, S61, S62, S65, S66, S68, S69, S70, S73, S79, S82, S84, S86, S87, S88.	Elicitation	28
S2, S6, S11, S31, S38, S40, S42, S43, S44, S46, S47, S49, S51, S52, S60, S77, S79, S81.	Analysis	18
S5, S6, S9, S15, S24 (2), S30, S37, S53, S56, S57 (2), S67, S80, S83, S89.	Other	16
S7, S10, S11, S19, S21, S26, S35, S46, S51, S54, S74, S79.	Validation	12
S1, S3, S50, S57, S71, S72.	Change Management	6

4.3. RQ3: Which Aspects of the Software Development Process, or of the Final Product, Were Affected by the Use of the CV?

Each primary study describes the impact by using different key terms, depending on the context in which the CV has been used. For that reason, during data extraction we obtained a broad set of terms related to this research question. In order to facilitate the analysis, we performed a thematic analysis following the guidelines of [48] and grouped all the original 154 impact mentions into 17 categories (listed in Table 10). The first column (Primary papers) shows the ID of the primary paper, the numbers in brackets indicate the quantity of mentions from the same primary paper. The column, Support Level, shows the quantity of impact mentions from our set of selected primary papers.

Table 10. Categories of Impacts.

Primary Paper	Category	Support Level
S02, S06, S08, S16, S20 (4), S23, S27, S32, S34, S40, S42, S45, S46, S50, S53 (2), S57, S59, S61, S62, S64, S65, S66, S68, S70, S75, S76 (2), S81, S86, S89 (2).	Guidance and understanding	35
S04, S10, S11, S12, S28, S33, S43, S47 (2), S51, S55, S56, S72, S77, S84, S85, S86, S89, S90 (2).	Automation and tool support	20
S02, S07, S11, S18, S19, S24, S26, S36, S39, S43, S44 (2), S46, S49 (2), S50, S84, S86.	Identification of conflicts and defects	18
S14, S17, S28, S29, S38, S67, S69, S74, S78, S82, S83, S84, S88 (2).	Completeness, correctness and accuracy	14
S11, S26, S35, S37, S73, S78, S79, S85, S86, S90.	Quality	10
S04, S05, S06, S11, S22, S30, S38, S52.	Reusability	8
S09, S10, S18, S41, S58, S70, S71.	Ambiguity	7
S02, S13, S25, S62, S68, S69, S87.	Elicitation	7
S21 (2), S28, S40, S49, S57, S63.	Testing and traceability	7
S07, S37, S60, S70, S79, S80.	Productivity and time reduction	6
S03, S31, S57, S65, S87.	Communication	5
S29, S38, S48, S54.	Consistency	4
S01 (2), S15, S40.	Evolution and maintainability	4
S37, S51, S72.	Costs	3
S12, S27, S74.	Validation	3
S03 (2)	Control and coordination	2
S77	Modelling	1

Guidance and understanding is the category that includes a greater number of mentions. The use of CVs, both ontologies and taxonomies, has a positive impact on the development of the software (process) and the final product (system) (S02). In the earlier phases of software development (i.e., requirements engineering) it is crucial to understand what data need to be requested from providers, where and when to capture the data, and what parts of them can be made accessible to different users (S20). Furthermore, a list of potential tasks (data manipulation) can help to overcome elicitation problems. CVs help requirements analysts in the process of understanding the key concepts in the problem space (S53). Problems of understanding typically result from the fact that RE involves a variety of people's expressed needs across different stakeholders. Another source of problems of understanding is the subjective nature of requirements, especially non-functional ones. The CVs provide a common terminology for relevant knowledge, enabling a uniform treatment of data and reducing the chances of missing out key requirements.

Automation and tool support is the second most cited category. CVs, especially ontologies, can be used as a theoretical foundation for automating some tasks, such as (S04): detecting conflicts, classify requirements or detect defects. Many tools, implementing ontologies and rules, provide support to detect NFR conflicts and inconsistencies (S43). Automatic analysis of natural language requirements and assistance to analysts during requirements specification is another benefit of using CVs. For example, the use of ontologies and boilerplates enables semi-automatic transformation of natural language requirements into boilerplate requirements (S90).

The identification of conflicts and defects is one of the most significant benefits of using CVs during RE. This benefit is closely related to automation and tool support. In this study, the term “conflict” refers to requirements that differ and may contradict each other. A good example of these conflicts due to legal cross-references can be found in (S49). Analyzing conflicts in software specifications is crucial when multiple stakeholder concerns need to be addressed (S44). The origin of conflicts can be inconsistencies, redundancies, overrides, and missing parts. A set of rules and the supporting CV (an ontology or taxonomy) can be stored and executed in knowledge-based systems to automatically detect conflicts. This ensures that requirements engineers do not overlook important compliance requirements.

When conflicts are identified and resolved, we achieve completeness, correctness and accuracy of requirements and their specifications. CVs have been widely used to improve completeness (S14), thanks to the ease they provide to verify that all the functionality requested by the stakeholders has been included in the requirements, and that all possible inputs (both valid and invalid) have been evaluated and have an adequate response (output). By using CVs and inference rules, an analyst can determine which requirements need to be added/removed to improve completeness (S29). Ontology-driven approaches seem to enable improved quality of requirements in terms of correctness, completeness and consistency (S78).

The impact of CVs on quality is addressed especially in 10 selected papers, but it is reported, indirectly, in most of the studies. The researcher used to relate quality to measurable criteria, such as: (a) completeness of individual requirement specifications; (b) existence of traceability relationships between requirements; (c) degree of compliance of individual requirements with the whole set of specifications (S37, S79). Some specific elements such as primary actors, conditions for actions, main flow of events, alternative flows, shall actions, object of action and destination of action, are measured to compute a quality level. Many of these elements are based on completeness, consistency and structure of the requirements.

Reusability is one of the most appreciated properties in software development. The CVs help increase reusability from two main perspectives: (a) facilitating the creation of reusable requirements (S05, S52) and (b) offering reusable mechanisms for the search, specification and validation of generic requirements (S22, S30). The application of CVs is frequent in CBSE (component-based software engineering) (S05) and SPLE (software product line engineering) (S30), providing greater structure to the planning and elicitation process, but also assisting in the process of configuration/adaptation of requirements from previous developments.

All other categories received a lower level of support since we did not find so many specific documents focused on them. However, given the degree of interrelation between the different categories, we cannot say that these categories are of minor importance. Many of the articles that mention aspects of productivity and time reduction are undoubtedly related to cost reduction. The improvements in consistency help the evolution of the requirements and facilitate their maintainability. The validation process is improved by increasing the quality of the communication, control and coordination processes of the development teams, thanks to the reduction in ambiguity.

One aspect to highlight is that we have not found a standardized use of the terms to describe the impacts. Some documents associate quality with terms closer to productivity (reduction in development time). We understand that it is difficult to classify the impacts, since there are many overlapping and causal relationships, for example, reducing ambiguity in the specifications increases the understanding of the requirements and, therefore, reduces the maintenance tasks, which, at the same time, saves effort (Productivity) and reduce Costs.

A second aspect is that all the papers reported positive effects of the use of the CVs. The most reported impacts for the requirements engineering process were a general increase in quality, thanks to the early identification of defects, inconsistencies and ambiguities; a general increase in productivity, due to the automation processes, which in turn are based on the support of specific tools.

The lack of enough reports about real-world applications is a frequent problem in software engineering. We believe that the large number of works included in this study allows us to draw conclusions with a solid empirical basis; however, we think that business–academy cooperation should be promoted, and the publication of more reports from the industry needs to be encouraged.

In the next two sections, we analyze the potential relationships when cross analyzing pairs of RQs.

4.4. Cross Analysis: RQ1 (CVs) vs. RQ2 (RE Activities)

This section compares and contrasts the results obtained from RQ1 and RQ2. Figure 3 shows the support level provided by ontologies and taxonomies to each of the RE activities. Table 11 shows the coverage of primary papers for every combination of CV and RE activity.

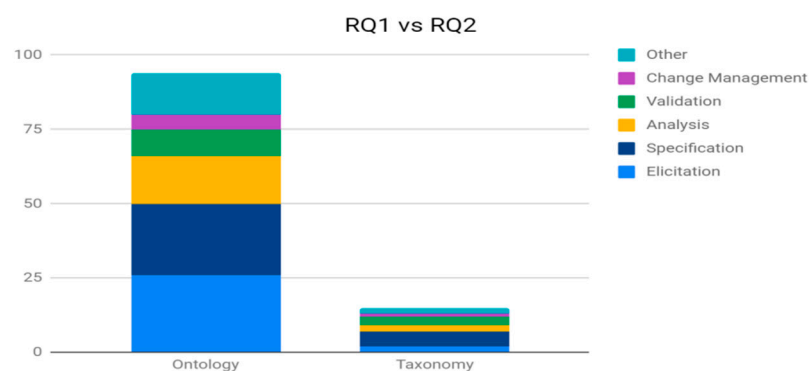


Figure 3. Research question (RQ)1 vs. RQ2.

Table 11. Coverage of RE activities in different CVs.

Type of CV	Elicitation	Specification	Analysis	Validation	Change Management	Other
Ontology	S02, S08, S12, S16, S27, S28, S29, S32, S34, S40, S55, S59, S61, S62, S65, S66, S68, S69, S70, S73, S79, S82, S84, S86, S87, S88	S04, S05, S17, S18, S22, S23, S25, S30, S33, S34, S36, S40, S41, S45, S48, S58, S63, S64, S65, S75, S76, S78, S85, S90	S02, S11, S31, S38, S40, S42, S43, S44, S46, S47, S51, S52, S60, S77, S79, S81	S10, S11, S19, S35, S46, S51, S54, S74, S79	S01, S03, S57, S71, S72	S05, S09, S15, S24 (2), S30, S37, S53, S56, S57 (2), S67, S80, S89
Taxonomy	S13, S39	S06, S14, S20, S26, S49	S06, S49	S07, S21, S26	S50	S06, S83

The cross analysis of RQ1 vs. RQ2 enables us to thoroughly assess the level of support given by CVs with respect to each stage of the RE lifecycle. The results of this comparative analysis depict convergence as Specification, Elicitation and Analysis of RE stages on average are still mostly supported by Ontologies and Taxonomies. Besides, the ratio of the support offered by the CVs to each RE stage is, in general, abundant.

Some of the key insights that are evident from Table 11 are: (1) Elicitation, Specification and Analysis stages of RE are considerably more supported by both ontologies and taxonomies so it is worth investigating how these two types of CVs specifically can be enhanced to bring more support to the remaining stages of RE lifecycle (Validation and Change Management). Nonetheless, the remaining

missing CVs (Folksonomies, Thesauri) also need to be investigated with the same aim to support the RE, and (2) investigate the suitability of combining CVs (hybrid approaches) and proposing automated tools to enhance the support level of these, less supported, RE tasks.

4.5. Cross Analysis: RQ3 vs. RQ2

We have analyzed the reported impacts (RQ3) and related them to the RE activities (RQ2) in which they occurred. Table 12 shows the support level (primary papers) identified for these relationships. Table 12 depicts the plotting of papers with RE activities whereas Figure 4 shows the overall trend of RE activities in respective papers.

Table 12. Categories of impacts in every RE activity.

RQ3/RQ2	Elicitation	Specification	Analysis	Validation	Change Mgmt.	Other
Guidance and understanding	S02, S08, S16, S27, S32, S34, S59, S61, S62, S66, S68, S70, S86	S06, S20(4), S23, S45, S64, S65, S75, S76(2)	S40, S42, S81	S46	S50, S57	S53(2), S89(2)
Automation and tool support	S12, S28, S55, S84, S86	S04, S33, S85, S90(2)	S11, S43, S47(2), S77	S10, S51	S72	S56, S89
Identification of conflicts and defects	S39, S84, S86	S18, S26, S36	S02, S43, S44(2), S49(2)	S07, S11, S19, S46	S50	S24
Completeness, correctness and accuracy	S28, S29, S69, S82, S84, S88(2)	S14, S17, S78	S38	S74		S67, S83
Quality	S73, S86	S26, S78, S85, S90		S11, S35, S79		S37
Reusability		S04, S22, S30	S06, S11, S38, S52			S05
Ambiguity	S70	S18, S41, S58		S10	S71	S09
Elicitation	S02, S13, S62, S68, S69, S87	S25				
Testing and traceability	S28	S63	S40, S49	S21(2)	S57	
Productivity and time reduction	S70		S60	S07, S79		S37, S80
Communication	S87	S65	S31		S03	S57
Consistency	S29	S48	S38	S54		
Evolution and maintainability	S40				S01(2)	S15
Costs			S51		S72	S37
Validation	S27, S12			S74		
Control and coordination					S03(2)	
Modelling			S77			

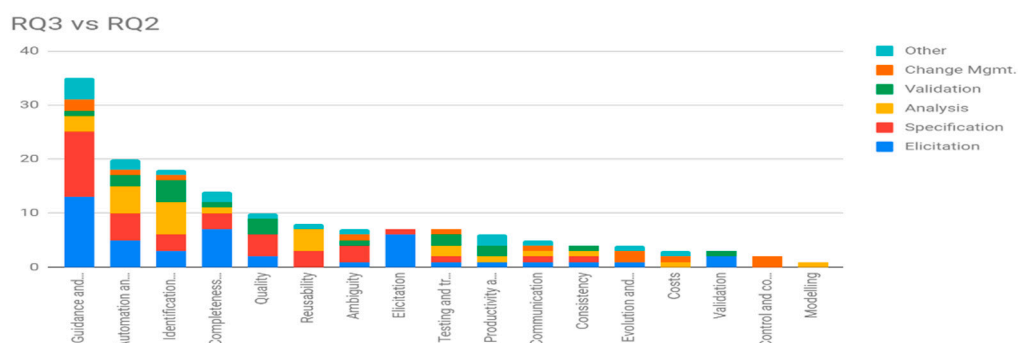


Figure 4. RQ3 vs. RQ2.

Figure 4 visualizes the distribution of references on each category of impact (RQ3) along the RE activities (RQ2). The activity with the highest support level is Elicitation (in blue color, next to the

horizontal axis) with 44 primary papers, in almost every category, except for Reusability, Costs, Control and Coordination, and Modelling. We think that the absence of impacts on these categories during the Elicitation is due to the early nature of the activity (first steps of RE). The vast majority of the support level is concentrated on Guidance and Understanding, by providing the developers with structured knowledge guides (ontologies and taxonomies) to conduct the elicitation process and facilitating the understanding of the collected requirements between all the stakeholders.

All the impacts of CVs affected the Specification activity, except Productivity and Time Reduction, Evolution and Maintainability, Costs, Validation, Control and Coordination and Modelling. Note that just because little or no support has been identified for a given RE activity, does not mean it does not exist. It is frequent that researchers report only those aspects that are more attractive or interesting, leaving aside the more trivial ones (least important).

Analysis is the third most mentioned RE activity, but far from the previous two. The reason for this decrease can be found in the work of disambiguation and specification that have been carried out in the first two activities. During the analysis of the requirements, they have been already refined enough to make the current use of the CVs not so critical. Figure 4 shows the details of this cross-analysis of the two variables, the Impacts and the RE activities where they happen.

4.6. DQ1: Most Active Researchers

A total of 243 researchers appear as authors in the 90 primary published research studies selected by our SMS study. We only report the top 10 active researchers, based on their number of published works, however, the full list can be accessed online as complementary material [44].

It is worth mentioning that the data allowed us to identify groups of researchers who often published collaborative works (e.g., Kossman and Odeh, Moser, Sindre and Stålhane, and Avdeenko and Murtazina). In addition, most of the active researchers are affiliated with academic organizations (80%). It is interesting to note that the top six active researchers, those with more than three publications, are based in Europe. Most of the identified researchers are currently active, except for the research group composed by Moser, Sindre and Stålhane, inactive since 2012. The top ten active researchers and their affiliations are shown in Table 13.

Table 13. Active researchers based on number of papers.

Author's Name	Organization	Support Level
Mario Kossman	AIRBUS, U.K.	4
Mohammed Odeh	Software Engineering Research Group, University of the West of England, U.K.	4
Thomas Moser	Institute of Software Technology and Interactive Systems, Vienna University of Technology, Austria	4
Guttorm Sindre	Department of Computer and Information Science, Norwegian University of Science and Technology, Norway	4
Tor Stålhane	Department of Computer and Information Science, Norwegian University of Science and Technology, Norway	4
Tatiana Avdeenko	Novosibirsk State Technical University, Russia	4
Chi-Lun Liu	Department of Multimedia and Mobile Commerce, Kainan University, Taiwan	3
Marina Murtazina	Novosibirsk State Technical University, Russia	3
Annie I. Antón	Department of Computer Science, North Carolina State University, USA	2
Kelly Antonini	University of the West of England, U.K.	2

4.7. DQ2: Most Active Organizations

We have identified 115 unique organizations based on the number of mentions in the published primary papers. Most of the identified organizations are academic institutions based in Europe. Table 14 depicts a list of the 12 most active organizations.

Table 14. Top 12 Active Organizations.

Organization's Name and Country	# Mentions
UIIT-PMAS Arid Agriculture University Rawalpindi, Pakistan	12
Norwegian University of Science and Technology, IDI/NTNU, Trondheim, Norway	11
Vienna University of Technology, Austria	10
Zhejiang University, Hangzhou, China	8
Novosibirsk State Technical University, Novosibirsk, Russia	8
CRI-Paris 1 Sorbonne University Paris, France	6
North Carolina State University, USA	6
University of Murcia, Spain	6
Concordia University, Montreal, Canada	5
Mälardalen University, Västerås, Sweden	5
Shinshu University, Nagano, Japan	5
University of Leeds, U.K.	5

It is noteworthy that there is an imbalance between the amount of research from academia and industry (82% academic, 11.3% industrial, 6.08% research institutes and 0.86% government agencies), which highlights the need for further research in industrial contexts, or in academia–industry collaborations.

4.8. DQ3: Most Active Countries

Based on the affiliation of all authors, we have been able to identify 37 different countries in the set of selected documents, located in the five continents. The three countries with the greatest representation are the United Kingdom, China and the United States. The rest of the countries do not reach a representation higher than 50% of the previous three. Figure 5 shows the number of authors whose affiliations belong to the ten countries with the highest number of mentions. The complete list of countries, showing a distribution of countries that offers a wide coverage of the five continents can be accessed as online complementary material [44].

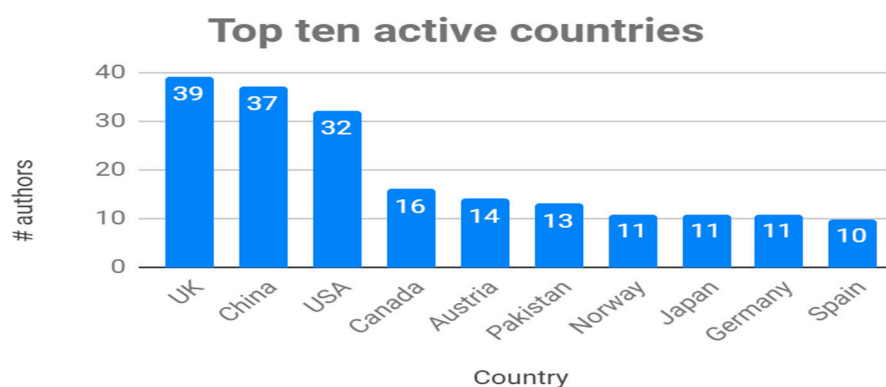


Figure 5. Most active countries (by authors' affiliation).

4.9. DQ4: Top Publishing Venues

The favorite venues for publication of the selected works were Conferences, with 64.4% of the primary works (Figure 6); the IEEE International Computer Software and Applications Conference, with four primary papers published, topped the list of the most cited (Table 15); together with the Workshops (10%), the Conferences cover 75% of the set of selected papers; however, 25.6% of the papers, published in Journals, is a good fact, and provides a solid empirical basis, given the quality and the recognized prestige of publication venues (Expert Systems, Requirements Engineering, etc.). Table 16 shows the complete list of mentioned journals.

Finally, Figure 7 shows that the interest in CVs and RE activities has remained at a moderate level (seven papers per year, on average) since 2009, with higher peaks in 2012 and 2016.

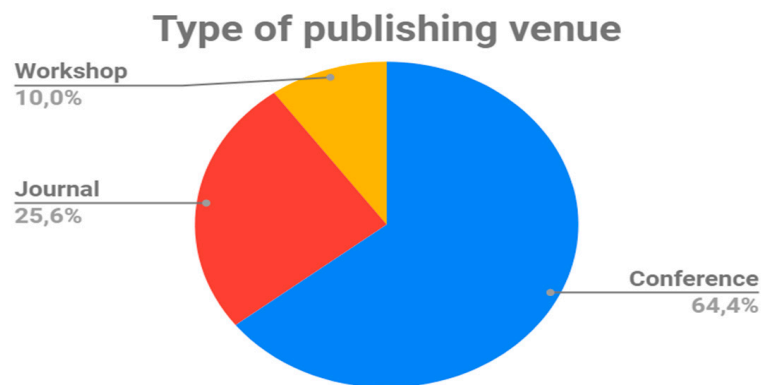


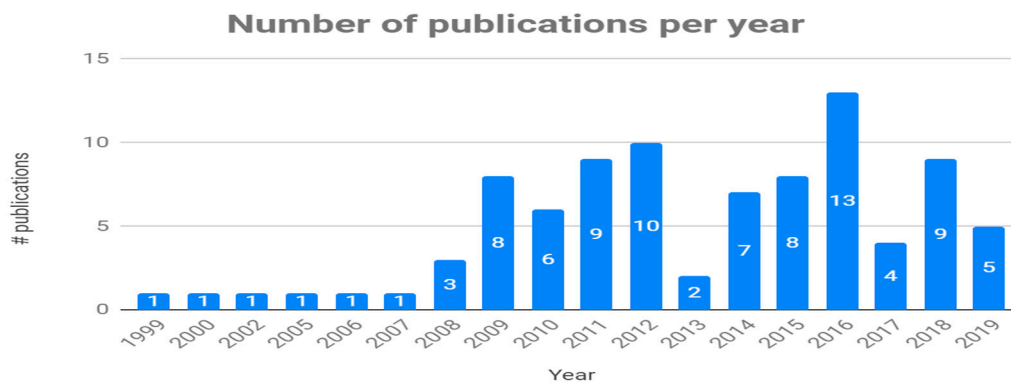
Figure 6. Distribution of publishing venues.

Table 15. Most cited conferences and workshops.

Conferences	Support Level
IEEE International Computer Software and Applications Conference	4
Asia-Pacific Software Engineering Conference, APSEC	3
IEEE International Conference on Global Software Engineering	2
International Conference on Information Integration and Web-based Applications and Services	2
INCOSE International Symposium	2
International Arab Conference on Information Technology (ACIT)	2
International Conference on Information Reuse and Integration (IRI)	2
International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)	2
SEKE	2
CEUR Workshop	2
IEEE International Conference on Data Mining Workshop	1
IEEE International Model-Driven Requirements Engineering Workshop (MoDRE)	1
IEEE International Requirements Engineering Conference Workshops, REW	1
IEEE International Workshop on Requirements Patterns, RePa	1
IEEE International Workshop on Semantic Computing and Systems	1
IEEE VIS Workshop on Evaluation and Beyond—Methodological Approaches for Visualization	1
International Workshop on Managing Requirements Knowledge, MARK	1

Table 16. Most cited journals.

Journal	Support Level
Expert Systems	2
Requirements Engineering	2
ACM Transactions on Management Information Systems (TMIS)	1
Artificial Intelligence for Engineering Design, Analysis and Manufacturing	1
Baltic Journal of Modern Computing	1
Computer Standards and Interfaces	1
IDDM (Oxford Journal)	1
IEEE Access	1
IEEE Software	1
IEICE TRANSACTIONS on Information and Systems	1
IET Software	1
Information and Software Technology	1
Information Systems Frontiers	1
International Journal of Human Computer Studies	1
International Journal of Intelligent Systems	1
International Journal of Software Engineering and Knowledge Engineering	1
ISRN Software Engineering	1
Journal of Convergence Information Technology	1
Journal of Research and Practice in Information Technology	1
Journal of Software	1
Procedia Computer Science	1

**Figure 7.** Evolution of the number of published papers.

5. Validity Threats

The five most frequently reported validity threats in software engineering research were considered [45,49]. In the next sections, details are provided about what we have done to avoid or mitigate these threats.

5.1. Descriptive Validity

Descriptive validity deals with threats to one's ability to capture and describe observations in an objective and precise way. To reduce this threat, work sessions using examples were conducted to achieve a common methodology for determining what data to extract and how they should be

extracted and a Data Extraction Form (DEF) was designed and agreed upon by all researchers [44]. Every entry in the DEF has an attached comment that links the value assigned by the researcher to specific text in the original source, increasing traceability and reducing researcher's bias.

5.2. Interpretive Validity

We have reduced this threat by applying two different mechanisms. First, five regular sessions were conducted (one session every two weeks) after the extraction data activity, to ensure that all the researchers agreed on the interpretation of the results (i.e., the raw data), a set of coding rules, and their implications. Second, authors, divided into two independent teams, derived the conclusions from the results. The first author compared and integrated the conclusions, homogenizing the writing style. In a final session, consisting of all the authors, the conclusions were double-checked to make sure that they could be linked, directly, to previous results stored in the DEF.

5.3. Theoretical Validity

To minimize the threat on the search process, we conducted two different search strategies: an automated search from four separate database sources and a complementary snowballing search process (backward and forward) from two seeds. The protocol to perform these two strategies is detailed in Section 3.2, Search strategies.

The details regarding the selection process and the protocol followed to reduce theoretical validity threats can be found in Section 3.3, Paper Selection: Inclusion/Exclusion Criteria.

5.4. Generalizability

To minimize the internal generalizability threat, we relied on the objectivity of the data extraction process and form (DEF), and the protocol to analyze the results and derive conclusions. Due to the sample size (90 primary papers), we think that the generalization of the results can be achieved with a low level of risk.

5.5. Reliability

To increase the reliability of this study, we performed a detailed report of the entire process, from the initial protocol to the conduction phase. Furthermore, the additional material accompanying this report provides, from our point of view, enough information to enable the replication of the study.

Finally, to reduce the threats to validity during the conducting phase of this study, we reported the rubrics used for the self-evaluation, following the guidelines from Petersen et al. [34]. These rubrics are available online, as complementary material, at [44].

6. Conclusions and Future Work

This study reports the planning and conducting phases of SMS. The SMS was based on the identification, analysis and classification of 90 selected published primary research papers, published from 1999 to 2019.

Our goal was to identify and classify the CVs, and to investigate their impact in RE activities. In addition, we also extracted a set of data related to demographic aspects, such as the most active researchers, organizations and countries in the area. The SMS provides a broad view of the research area for the last 20 years.

We selected 90 studies from 2348 papers retrieved by two complementary search strategies. The extracted data allow us to conclude that:

- **RQ1:** Approximately 88% of the selected studies reported the use of ontologies and 12% focused on taxonomies. There is a lack of direct empirical evidence on the use of folksonomies or thesauri, although some studies reported their application, but were always embedded, as part of ontology or a thesaurus.

- **RQ2:** The use of CVs has been studied in all RE activities, but not with the same interest. The applications have focused on the activities of elicitation (31%) and specification (32%), whereas the empirical support was reduced for the activities of analysis (20%), validation (13%) and change management (7%).
- **RQ3:** The impacts with greater empirical support have been: guidance and understanding (39%), automation and support of tools (22%), and help to identify conflicts and defects (20%). Many of these impacts have overlapping effects, for example, by facilitating the understanding of the requirements, conflicts and possible defects are reduced, increasing the quality of the final specifications.

Demographic data show that interest in CVs and their application in RE activities remain high in the research community. Several research groups, mainly in Europe, have been active since 2010, with an average, in the last 5 years, of eight articles published per year. Although most authors belong to academic organizations, there are also representatives from industry, government organizations and research institutes. A large part of the research is carried out in European organizations. The results highlight Pakistan as a very active country, with 12 mentions, however, its real participation is only two works (S54 and S71), with seven authors in the first study and five in the second (three researchers participated in both studies). On the other hand, the data show that the favorite publishing venues are Conferences and Workshops (almost 75% of the cases), which would suggest a greater participation of authors with industrial affiliations, however, their presence, based on the set of selected papers, is low (around 20%). We aim for the community to publish more collaborative works (academy–industry), to overcome the lack of industrial empirical evidence.

As future work, we suggest that:

- We believe it might be interesting to investigate the use of CVs in development environments based on open source models, where RE activities involve leveraging online comments and the wisdom of the crowd.
- There is a need for more empirical research by conducting comprehensive review on how ontologies support the whole software engineering process.
- In this SMS study, we did not find direct evidence on exploring the suitability and impact of folksonomies and thesauri on RE so it might be good to be investigated.
- More collaborative empirical research needs to be conducted; more specifically, industrial–academic collaborations to evaluate the suitability of different CVs.
- To conduct evaluation studies that can compare different RE processes supported by different types of controlled vocabularies used.

Author Contributions: Conceptualization, A.A. and J.L.B.J.; methodology, A.A., J.L.B.J. and C.F.; validation, A.A., J.L.B.J. and C.F.; formal analysis, A.A., J.L.B.J. and C.F.; data curation, A.A., and J.L.B.J.; writing—original draft preparation, A.A., and J.L.B.J.; writing—review and editing, A.A., J.L.B.J., A.A.K. and C.F.; supervision, J.L.B.J. and C.F.; project administration, J.L.B.J. and C.F.; funding acquisition, C.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work and the APC were funded by the National Key R&D Program of China (No. 2017YFB1002101), National Natural Science Foundation of China (No. U1636203), and the Joint Advanced Research Foundation of China Electronics Technology Group Corporation (CETC) (No. 6141B08010102).

Acknowledgments: This work was supported by the National Key R&D Program of China (No. 2017YFB1002101), National Natural Science Foundation of China (No. U1636203), and the Joint Advanced Research Foundation of China Electronics Technology Group Corporation (CETC) (No. 6141B08010102).

Conflicts of Interest: All the authors declare that they have no conflict of interests relevant to this article.

Appendix A

Table A1. Selected papers.

Paper ID	Full Bibliographic Reference
S01	Abbasipour, M., Sackmann, M., Khendek, F., & Toeroe, M. (2014, August). Ontology-based user requirements decomposition for component selection for highly available systems. In <i>Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014)</i> (pp. 44–51). IEEE.
S02	Al Balushi, T. H., Sampaio, P. R. F., & Loucopoulos, P. (2013). Eliciting and prioritizing quality requirements supported by ontologies: a case study using the Elicit O framework and tool. <i>Expert Systems</i> , 30(2), 129–151.
S03	Alsanad, A. A., Chikh, A., & Mirza, A. (2019). A Domain Ontology for Software Requirements Change Management in Global Software Development Environment. <i>IEEE Access</i> , 7, 49352–49361.
S04	Amri, S. K., Darmoul, S., & Hajri-Gabouj, S. (2018, March). An immune designed ontology to specify requirements and recommend layout solutions. In <i>2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)</i> (pp. 255–260). IEEE.
S05	Andreou, A. S., & Papatheocharous, E. (2015, April). Automatic matching of software component requirements using semi-formal specifications and a CBSE ontology. In <i>2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)</i> (pp. 118–128). IEEE.
S06	Antón, A. I., Earp, J. B., & Reese, A. (2002, September). Analyzing website privacy requirements using a privacy goal taxonomy. In <i>Proceedings IEEE Joint International Conference on Requirements Engineering</i> (pp. 23–31). IEEE.
S07	Anu, V. K., Walia, G. S., Hu, W., Carver, J., & Bradshaw, G. L. (2016, July). Effectiveness of Human Error Taxonomy during Requirements Inspection: An Empirical Investigation. In <i>SEKE</i> (pp. 531–536).
S08	Aranda, G. N., Vizcaíno, A., & Piattini, M. (2009, July). Analyzing ontology as a facilitator during global requirements elicitation. In <i>2009 Fourth IEEE International Conference on Global Software Engineering</i> (pp. 309–314). IEEE.
S09	Assawamekin, N. (2011). Resolving semantic heterogeneity in multiperspective requirements traceability using ontology matching. <i>Journal of Convergence Information Technology</i> , Vol. 6, No. 6 (June 2011).
S10	Avdeenko, T. V., & Pustovalova, N. V. (2016, October). The ontology-based approach to support the requirements engineering process. In <i>2016 13th International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)</i> (Vol. 2, pp. 513–518). IEEE.
S11	Bagriyanik, S., & Karahoca, A. (2016). Automated COSMIC Function Point measurement using a requirements engineering ontology. <i>Information and Software Technology</i> , 72, 189–203.
S12	Bargui, F., Ben-Abdallah, H., & Feki, J. (2011, November). A decision making ontology building process for analytical requirements elicitation. In <i>2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications</i> (pp. 1529–1536). IEEE.
S13	Burnay, C. (2016). Are stakeholders the only source of information for requirements engineers? Toward a taxonomy of elicitation information sources. <i>ACM Transactions on Management Information Systems (TMIS)</i> , 7(3), 8.
S14	Chance, B. D., & Melhart, B. E. (1999, March). A taxonomy for scenario use in requirements elicitation and analysis of software systems. In <i>Proceedings ECBS'99. IEEE Conference and Workshop on Engineering of Computer-Based Systems</i> (pp. 232–238). IEEE.
S15	Chen, F., Zhou, H., Yang, H., Ward, M., & Chu, W. C. C. (2011, July). Requirements recovery by matching domain ontology and program ontology. In <i>2011 IEEE 35th Annual Computer Software and Applications Conference</i> (pp. 602–607). IEEE.
S16	Chen, X., & Jin, Z. (2016). Capturing requirements from expected interactions between software and its interactive environment: an ontology based approach. <i>International Journal of Software Engineering and Knowledge Engineering</i> , 26(01), 15–39.

Table A1. Cont.

Paper ID	Full Bibliographic Reference
S17	Chen, X., Ye, R., Sun, H., & Lu, H. (2013, July). Deriving Requirements Specification with Time: A Software Environment Ontology Based Approach. In <i>2013 IEEE 37th Annual Computer Software and Applications Conference</i> (pp. 431–436). IEEE.
S18	Daramola, O., Sindre, G., & Stalhane, T. (2012, September). Pattern-based security requirements specification using ontologies and boilerplates. In <i>2012 Second IEEE international workshop on requirements patterns (RePa)</i> (pp. 54–59). IEEE.
S19	Djilani, Z., Khat, A., Khouri, S., & Bellatreche, L. (2016, November). Murgroom: multi-site requirement reuse through graph and ontology matching. In <i>Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services</i> (pp. 160–169). ACM.
S20	Elshehal, M., Alvarado, N., McVey, L., Randell, R., Mamas, M., & Ruddie, R. A. (2018, October). From Taxonomy to Requirements: A Task Space Partitioning Approach. In <i>Proceedings of the IEEE VIS Workshop on Evaluation and Beyond—Methodological Approaches for Visualization (BELIV)</i> . IEEE.
S21	Felderer, M., & Beer, A. (2014). Using defect taxonomies for testing requirements. <i>IEEE Software</i> , 32(3), 94–101.
S22	Greenwell, R., Liu, X., Chalmers, K., & Pahl, C. (2016). A Task Orientated Requirements Ontology for Cloud Computing Services. SciTePress.
S23	Hasan, M. M., Aganostopoulos, D., Loucopoulos, P., & Nikolaidou, M. (2017, March). Regulatory Requirements Compliance in e-Government System Development: an Ontology Framework. In <i>Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance</i> (pp. 441–449). ACM.
S24	He, H., Wang, Z., Zhang, Y., & Zhang, W. (2012). An Ontology-Based Framework of Requirements Evolvment Management. <i>JSW</i> , 7(9), 2018–2025.
S25	Hovorushchenko, T., Pavlova, O., & Fedula, M. (2018). Improving the Input Information for Medical Software Requirements Specifications using Ontology-Based Intelligent Agent. In <i>IDDM</i> (pp. 113–125).
S26	Hu, W., Carver, J. C., Anu, V. K., Walia, G. S., & Bradshaw, G. (2016, September). Detection of requirement errors and faults via a human error taxonomy: a feasibility study. In <i>Proceedings of the 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement</i> (p. 30). ACM.
S27	Jiang, H., & Yang, X. (2009, January). Performance requirement elicitation for financial information system based on ontology. In <i>TENCON 2009–2009 IEEE Region 10 Conference</i> (pp. 1–5). IEEE.
S28	Jingbai, T., Keqing, H., Chong, W., & Wei, L. (2008, July). A context awareness non-functional requirements metamodel based on domain ontology. In <i>IEEE International Workshop on Semantic Computing and Systems</i> (pp. 1–7). IEEE.
S29	Kaiya, H., & Saeki, M. (2006, September). Using domain ontology as domain knowledge for requirements elicitation. In <i>14th IEEE International Requirements Engineering Conference (RE'06)</i> (pp. 189–198). IEEE.
S30	Karatas, E. K., Iyidir, B., & Birtürk, A. (2014, December). Ontology-based software requirements reuse: Case study in fire control software product line domain. In <i>2014 IEEE International Conference on Data Mining Workshop</i> (pp. 832–839). IEEE.
S31	Kassab, M., Ormandjieva, O., & Daneva, M. (2009, September). An ontology based approach to non-functional requirements conceptualization. In <i>2009 Fourth International Conference on Software Engineering Advances</i> (pp. 299–308). IEEE.
S32	Kifle, M. (2012, September). Multi-perspective Ontology to Understand Organizational Requirements. In <i>2012 African Conference for Software Engineering and Applied Computing</i> (pp. 67–74). IEEE.
S33	Kluge, R., Hering, T., Belter, R., & Franczyk, B. (2008, July). An approach for matching functional business requirements to standard application software packages via ontology. In <i>2008 32nd Annual IEEE International Computer Software and Applications Conference</i> (pp. 1017–1022). IEEE.

Table A1. Cont.

Paper ID	Full Bibliographic Reference
S34	Koay, N., Kataria, P., Juric, R., Oberndorf, P., & Terstyanszky, G. (2009, January). Ontological support for managing non-functional requirements in pervasive healthcare. In <i>2009 42nd Hawaii International Conference on System Sciences</i> (pp. 1–10). IEEE.
S35	Kof, L., Gacitua, R., Rouncefield, M., & Sawyer, P. (2010, September). Ontology and model alignment as a means for requirements validation. In <i>2010 IEEE Fourth International Conference on Semantic Computing</i> (pp. 46–51). IEEE.
S36	Körner, S. J., & Brumm, T. (2009, July). Improving Natural Language Specifications with Ontologies. In <i>SEKE</i> (pp. 552–557).
S37	Kossmann, M., & Odeh, M. (2010, July). 7.4. 3 Ontology-driven Requirements Engineering—A case study of OntoREM in the aerospace context. In <i>INCOSE International Symposium</i> (Vol. 20, No. 1, pp. 1000–1012).
S38	Velasco, J. L., Valencia-García, R., Fernández-Breis, J. T., & Toval, A. (2009). Modelling reusable security requirements based on an ontology framework. <i>Journal of Research and Practice in Information Technology</i> , 41(2), 119.
S39	Laurent, P., Mader, P., Cleland-Huang, J., & Steele, A. (2010, August). A taxonomy and visual notation for modeling globally distributed requirements engineering projects. In <i>2010 5th IEEE International Conference on Global Software Engineering</i> (pp. 35–44). IEEE.
S40	Lee, S. W., & Gandhi, R. A. (2005, December). Ontology-based active requirements engineering framework. In <i>12th Asia-Pacific Software Engineering Conference (APSEC'05)</i> (pp. 8–pp). IEEE.
S41	Lemke, M. T., Stone, R. B., & Arlitt, R. A. (2017, August). Ontologies to Support Customer Requirement Formulation in Aerospace Design. In <i>ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference</i> (pp. V007T06A018–V007T06A018). American Society of Mechanical Engineers.
S42	Lin, C. Y. I., & Ho, C. S. (2000). Generating domain knowledge for requirement analysis based on acquisition ontology. <i>International journal of intelligent systems</i> , 15(12), 1125–1155.
S43	Liu, C. L. (2016). CDNFRE: Conflict detector in non-functional requirement evolution based on ontologies. <i>Computer Standards & Interfaces</i> , 47, 62–76.
S44	Liu, C. L., & Huang, H. H. (2015). Ontology-Based Requirement Conflicts Analysis in Class Diagrams. In <i>Proceedings of the World Congress on Engineering</i> (Vol. 1).
S45	Lopata, A., & Makrickienė, N. Requirements Engineering, Supported by Ontology and Enterprise Modelling.
S46	Lopez-Lorca, A. A., Beydoun, G., Valencia-Garcia, R., & Martinez-Bejar, R. (2016). Supporting agent oriented requirement analysis with ontologies. <i>International Journal of Human-Computer Studies</i> , 87, 20–37.
S47	Murtazina, M. S., & Avdeenko, T. V. (2018, October). Ontology-Based Approach to the Requirements Engineering in Agile Environment. In <i>2018 XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)</i> (pp. 496–501). IEEE.
S48	Mahmud, N., Seceleanu, C., & Ljungkrantz, O. (2015, June). ReSA: An ontology-based requirement specification language tailored to automotive systems. In <i>10th IEEE International Symposium on Industrial Embedded Systems (SIES)</i> (pp. 1–10). IEEE.
S49	Maxwell, J. C., Antón, A. I., Swire, P., Riaz, M., & McCraw, C. M. (2012). A legal cross-references taxonomy for reasoning about compliance requirements. <i>Requirements Engineering</i> , 17(2), 99–115.
S50	McGee, S., & Greer, D. (2011, August). Software requirements change taxonomy: Evaluation by case study. In <i>2011 IEEE 19th International Requirements Engineering Conference</i> (pp. 25–34). IEEE.
S51	Moshirpour, M., Mireslami, S., Alhaji, R., & Far, B. H. (2012, August). Automated ontology construction from scenario based software requirements using clustering techniques. In <i>2012 IEEE 13th International Conference on Information Reuse & Integration (IRI)</i> (pp. 541–547). IEEE.
S52	Mukhopadhyay, A., & Ameri, F. (2016). An ontological approach to engineering requirement representation and analysis. <i>AI EDAM</i> , 30(4), 337–352.

Table A1. Cont.

Paper ID	Full Bibliographic Reference
S53	Murtazina, M. S., & Avdeenko, T. V. (2019). An Ontology-based Approach to Support for Requirements Traceability in Agile Development. <i>Procedia Computer Science</i> , 150, 628–635.
S54	Nazir, S., Motla, Y. H., Abbas, T., Khatoon, A., Jabeen, J., Iqra, M., & Bakhat, K. (2014, November). A process improvement in requirement verification and validation using ontology. In <i>Asia-Pacific World Congress on Computer Science and Engineering</i> (pp. 1–8). IEEE.
S55	Omoronyia, I., Sindre, G., Stålhane, T., Biffl, S., Moser, T., & Sunindyo, W. (2010, June). A domain ontology building process for guiding requirements elicitation. In <i>International working conference on requirements engineering: Foundation for software quality</i> (pp. 188–202). Springer, Berlin, Heidelberg.
S56	Pakdeetrakulwong, U., Wongthongtham, P., & Khan, N. (2015, September). An Ontology-Based Multi-Agent System to Support Requirements Traceability in Multi-Site Software Development Environment. In <i>Proceedings of the ASWEC 2015 24th Australasian Software Engineering Conference</i> (pp. 96–100). ACM.
S57	Pires, P. F., Delicato, F. C., Cobe, R., Batista, T., Davis, J. G., & Song, J. H. (2011). Integrating ontologies, model driven, and CNL in a multi-viewed approach for requirements engineering. <i>Requirements Engineering</i> , 16(2), 133–160.
S58	Polpinij, J. (2009, December). An ontology-based text processing approach for simplifying ambiguity of requirement specifications. In <i>2009 IEEE Asia-Pacific Services Computing Conference (APSCC)</i> (pp. 219–226). IEEE.
S59	Provenzano, L., Hänninen, K., Zhou, J., & Lundqvist, K. (2017, December). An Ontological Approach to Elicit Safety Requirements. In <i>2017 24th Asia-Pacific Software Engineering Conference (APSEC)</i> (pp. 713–718). IEEE.
S60	Roh, W., & Lee, S. W. (2017, September). An Ontological Approach to Predict Trade-Offs between Security and Usability for Mobile Application Requirements Engineering. In <i>2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)</i> (pp. 69–75). IEEE.
S61	Salinesi, C., Ivankina, E., & Angole, W. (2008, September). Using the RITA threats ontology to guide requirements elicitation: an empirical experiment in the banking sector. In <i>2008 First International Workshop on Managing Requirements Knowledge</i> (pp. 11–15). IEEE.
S62	Salini, P., & Kanmani, S. (2016, February). A novel method: Ontology-based security requirements engineering framework. In <i>2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)</i> (pp. 1–5). IEEE.
S63	Murtazina, M. S., & Avdeenko, T. V. (2018). The ontology-driven approach to support the requirements engineering process in Scrum framework.
S64	Shah, U., Patel, S., & Jinwala, D. (2019, March). An Ontological Approach to Specify Conflicts among Non-Functional Requirements. In <i>Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis</i> (pp. 145–149). ACM.
S65	Sithithanasakul, S., & Choosri, N. (2016, December). Using ontology to enhance requirement engineering in agile software process. In <i>2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)</i> (pp. 181–186). IEEE.
S66	Souag, A., Salinesi, C., Mazo, R., & Comyn-Wattiau, I. (2015, March). A security ontology for security requirements elicitation. In <i>International symposium on engineering secure software and systems</i> (pp. 157–177). Springer, Cham.
S67	Umoh, E., Sampaio, P. R. F., & Theodoulidis, B. (2011, July). REFINTO: An ontology-based requirements engineering framework for business-IT alignment in financial services organizations. In <i>2011 IEEE International Conference on Services Computing</i> (pp. 600–607). IEEE.
S68	Wang, T., Si, Y., Xuan, X., Wang, X., Yang, X., Li, S., & Kavs, A. J. (2010, November). A QoS ontology cooperated with feature models for non-functional requirements elicitation. In <i>Proceedings of the Second Asia-Pacific Symposium on Internetware</i> (p. 17). ACM.
S69	Xiang, J., Liu, L., Qiao, W., & Yang, J. (2007, July). Srem: A service requirements elicitation mechanism based on ontology. In <i>31st Annual International Computer Software and Applications Conference (COMPSAC 2007)</i> (Vol. 1, pp. 196–203). IEEE.

Table A1. Cont.

Paper ID	Full Bibliographic Reference
S70	Yuan, X., & Tripathi, S. (2016, August). An approach of dynamically combining ontologies for interactive Requirements Elicitation. In <i>2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)</i> (pp. 11–16). IEEE.
S71	Khatoun, A., Motla, Y. H., Azeem, M., Naz, H., & Nazir, S. (2013, December). Requirement change management for global software development using ontology. In <i>2013 IEEE 9th International Conference on Emerging Technologies (ICET)</i> (pp. 1–6). IEEE.
S72	Liu, C. L., & Yang, H. L. (2012). Applying ontology-based blog to detect information system post-development change requests conflicts. <i>Information Systems Frontiers</i> , 14(5), 1019–1032.
S73	Dzung, D. V., & Ohnishi, A. (2009, August). Improvement of quality of software requirements with requirements ontology. In <i>2009 Ninth International Conference on Quality Software</i> (pp. 284–289). IEEE.
S74	Dzung, D. V., & Ohnishi, A. (2014). Ontology-Based Checking Method of Requirements Specification. <i>IEICE TRANSACTIONS on Information and Systems</i> , 97(5), 1028–1038.
S75	Barra, E., & Morato, J. (2015, June). Knowledge Identification from Requirements Specification. In <i>International Symposium on Languages, Applications and Technologies</i> (pp. 264–270). Springer, Cham.
S76	Qaddoumi, E., Odeh, M., Khan, Z., & Kossmann, M. (2018, November). OntoSoS. QR: Semantic representation of quality requirements metamodel for systems of systems. In <i>2018 International Arab Conference on Information Technology (ACIT)</i> (pp. 1–7). IEEE.
S77	Hu, H., Ma, Q., Zhang, T., Tan, Y., Xiang, H., Fu, C., & Feng, Y. (2015). Semantic modelling and automated reasoning of non-functional requirement conflicts in the context of softgoal interdependencies. <i>IET Software</i> , 9(6), 145–156.
S78	Antonini, K., Lange, C., Kossmann, M., & Odeh, M. (2014, October). Applying OntoREM to a space mission: developing requirements for the MASCOT lander. In <i>INCOSE International Symposium</i> (Vol. 24, No. s1, pp. 1–12).
S79	Antonini, K., Odeh, M., Kossmann, M., & Lange, C. (2014, December). Evaluating the effectiveness of mindmapping in generating domain ontologies using OntoREM: the MASCOT case study. In <i>Proc. of the International Arab Conference on Information Technology, Nizwa, Oman</i> (pp. 247–255).
S80	Bhushan, M., Goel, S., & Kumar, A. (2018). Improving quality of software product line by analysing inconsistencies in feature models using an ontological rule-based approach. <i>Expert Systems</i> , 35(3), e12256.
S81	Kost, M., & Freytag, J. C. (2012, February). Privacy analysis using ontologies. In <i>Proceedings of the second ACM conference on Data and Application Security and Privacy</i> (pp. 205–216). ACM.
S82	Farfeleder, S., Moser, T., Krall, A., Stålhane, T., Omoronyia, I., & Zojer, H. (2011, May). Ontology-driven guidance for requirements elicitation. In <i>Extended Semantic Web Conference</i> (pp. 212–226). Springer, Berlin, Heidelberg.
S83	Gopalakrishnan, S., Karpati, P., & Sindre, G. (2011). Evaluating a Taxonomy for Mobility Requirements by a Controlled Experiment. <i>ISRN Software Engineering</i> , 2012.
S84	Yuan, X., & Zhang, X. (2015, August). An ontology-based requirement modeling for interactive software customization. In <i>2015 IEEE International Model-Driven Requirements Engineering Workshop (MoDRE)</i> (pp. 1–10). IEEE.
S85	Farfeleder, S., Moser, T., Krall, A., Stålhane, T., Zojer, H., & Panis, C. (2011, April). DODT: Increasing requirements formalism using domain ontologies for improved embedded systems development. In <i>14th IEEE International Symposium on Design and Diagnostics of Electronic Circuits and Systems</i> (pp. 271–274). IEEE.
S86	Veleda, R., & Cysneiros, L. M. (2019, June). Towards an Ontology-Based Approach for Eliciting Possible Solutions to Non-Functional Requirements. In <i>International Conference on Advanced Information Systems Engineering</i> (pp. 145–161). Springer, Cham.
S87	Li, G., Jin, Z., Xu, Y., & Lu, Y. (2011, December). An engineerable ontology based approach for requirements elicitation in process centered problem domain. In <i>International Conference on Knowledge Science, Engineering and Management</i> (pp. 208–220). Springer, Berlin, Heidelberg.

Table A1. Cont.

Paper ID	Full Bibliographic Reference
S88	Kaiya, H., Shimizu, Y., Yasui, H., Kaijiri, K., & Saeki, M. (2010, November). Enhancing domain knowledge for requirements elicitation with web mining. In <i>2010 Asia Pacific Software Engineering Conference</i> (pp. 3–12). IEEE.
S89	Makrickienė, N., Gudas, S., & Lopata, A. (2019). Ontology and enterprise modelling driven software requirements development approach. <i>Baltic journal of modern computing</i> , 190–210.
S90	Daramola, O., Sindre, G., & Moser, T. (2012, September). Ontology-based support for security requirements specification process. In <i>OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"</i> (pp. 194–206). Springer, Berlin, Heidelberg.

References

- Dick, J.; Hull, E.; Jackson, K. *Requirements Engineering*; Springer International Publishing: Cham, Switzerland, 2017.
- Leise, F. Controlled vocabularies: An introduction. *Index* **2008**, *26*, 121–126. [[CrossRef](#)]
- Pizard, S.; Vallespir, D. Towards a controlled vocabulary on software engineering education. *Eur. J. Eng. Educ.* **2017**, *42*, 927–943. [[CrossRef](#)]
- NISO. ANSI/NISO Z39.19-2005: Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. NISO. 2010. Available online: <https://www.niso.org/publications/ansiniso-z3919-2005-r2010> (accessed on 30 July 2019).
- Robertson, S.; Robertson, J. *Mastering the Requirements Process: Getting Requirements Right*; Addison-wesley: Columbia, CA, USA, 2012.
- Kotonya, G.; Sommerville, I. *Requirements Engineering: Processes and Techniques*; Wiley Publishing: Columbia, CA, USA, 1998.
- Parreira, P.A.; Penteado, R.A.D. Domain ontologies in the context of Requirements Engineering. In Proceedings of the 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA), Marrakech, Morocco, 17–20 November 2015; pp. 1–8.
- Dermeval, D.; Vilela, J.; Bittencourt, I.I.; Castro, J.; Isotani, S.; Brito, P.; Silva, A. Applications of ontologies in requirements engineering: A systematic review of the literature. *Requir. Eng.* **2016**, *21*, 405–437. [[CrossRef](#)]
- Anu, V.; Hu, W.; Carver, J.C.; Walia, G.S.; Bradshaw, G. Development of a human error taxonomy for software requirements: A systematic literature review. *Inf. Softw. Technol.* **2018**, *103*, 112–124. [[CrossRef](#)]
- Jayatilleke, S.; Lai, R. A systematic review of requirements change management. *Inf. Softw. Technol.* **2018**, *93*, 163–185. [[CrossRef](#)]
- Pacheco, C.; García, I.; Reyes, M. Requirements elicitation techniques: A systematic literature review based on the maturity of the techniques. *IET Softw.* **2018**, *12*, 365–378. [[CrossRef](#)]
- Budgen, D.; Turner, M.; Brereton, P.; Kitchenham, B.A. Using Mapping Studies in Software Engineering. *PPIG* **2008**, *2*, 195–204.
- Kitchenham, B.A.; Budgen, D.; Brereton, O.P. Using mapping studies as the basis for further research—A participant-observer case study. *Inf. Softw. Technol.* **2011**, *53*, 638–651. [[CrossRef](#)]
- Petersen, K.; Feldt, R.; Mujtaba, S.; Mattsson, M. Systematic mapping studies in software engineering. In Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), Bari BA, Italy, 26–28 June 2008; pp. 68–77.
- Alencar, G.A.; Felipe, V.D.S.; Correia-Neto, J.D.S.; Teixeira, M.M. Non-Functional Requirements in Health Information Systems. In Proceedings of the 2019 14th Iberian Conference on Information Systems and Technologies (CISTI), Coimbra, Portugal, 19–22 June 2019; pp. 1–5.
- Horkoff, J.; Aydemir, F.B.; Cardoso, E.; Li, T.; Maté, A.; Paja, E.; Salnitri, M.; Piras, L.; Mylopoulos, J.; Giorgini, P. Goal-oriented requirements engineering: An extended systematic mapping study. *Requir. Eng.* **2019**, *24*, 133–160. [[CrossRef](#)] [[PubMed](#)]
- Souza, E.; Moreira, A.; Goulão, M. Deriving architectural models from requirements specifications: A systematic mapping study. *Inf. Softw. Technol.* **2019**, *109*, 26–39. [[CrossRef](#)]

18. Barros-Justo, J.L.; Benitti, F.B.V.; Cravero-Leal, A.L. Software patterns and requirements engineering activities in real-world settings: A systematic mapping study. *Comput. Stand. Interfaces* **2018**, *58*, 23–42. [\[CrossRef\]](#)
19. Curcio, K.; Navarro, T.; Malucelli, A.; Reinehr, S. Requirements engineering: A systematic mapping study in agile software development. *J. Syst. Softw.* **2018**, *139*, 32–50. [\[CrossRef\]](#)
20. Rehman, S.; Gruhn, V.; Shafiq, S.; Inayat, I. A Systematic Mapping Study on Security Requirements Engineering Frameworks for Cyber-Physical Systems. In Proceedings of the International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage, Melbourne, NSW, Australia, 11–13 December 2018; pp. 428–442.
21. Vegendla, A.; Duc, A.N.; Gao, S.; Sindre, G. A systematic mapping study on requirements engineering in software ecosystems. *J. Inf. Technol. Res. (JITR)* **2018**, *11*, 49–69. [\[CrossRef\]](#)
22. Siegemund, K.; Thomas, E.J.; Zhao, Y.; Pan, J.; Assmann, U. Towards ontology-driven requirements engineering. In Proceedings of the Workshop Semantic Web Enabled Software Engineering at 10th International Semantic Web Conference (ISWC), Bonn, Germany, 23–27 October 2011.
23. Fatwanto, A. Software requirements specification analysis using natural language processing technique. In Proceedings of the 2013 International Conference on QiR, Yogyakarta, Indonesia, 25–28 June 2013; pp. 105–110.
24. Bianchini, D. Deriving Folksonomies for Improving Web API Search. In Proceedings of the OTM Confederated International Conferences on the Move to Meaningful Internet Systems, Amantea, Italy, 27–31 October 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 517–534.
25. Barros-Justo, J.L. Mining unstructured data to support requirements elicitation by using controlled vocabularies: A systematic mapping study. *Dyna* **2015**, *82*, 165–169. [\[CrossRef\]](#)
26. Polpinij, J.; Ghose, A. An automatic elaborate requirement specification by using hierarchical text classification. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Hubei, China, 12–14 December 2008; pp. 706–709.
27. Merten, T.; Schäfer, T.; Bürsner, S. Using RE knowledge to assist automatically during requirement specification. In Proceedings of the 2012 Seventh IEEE International Workshop on Requirements Engineering Education and Training (REET), Chicago, IL, USA, 24 September 2012; pp. 9–13.
28. Medelyan, O.; Witten, I.H.; Divoli, A.; Broekstra, J. Automatic construction of lexicons, taxonomies, ontologies, and other knowledge structures. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2013**, *3*, 257–279. [\[CrossRef\]](#)
29. Hedden, H. Controlled vocabularies, thesauri, and taxonomies. *Index. Int. J. Index.* **2008**, *26*, 33–34. [\[CrossRef\]](#)
30. Sommerville, I. *Software Engineering*, 10th ed.; Pearson Higher Education: London, UK, 2016.
31. IEEE/ISO/IEC. *Systems and Software Engineering—Life Cycle Processes—Requirements Engineering (ISO/IEC/IEEE 29148)*; International Standard Organization (ISO): Geneva, Switzerland, 2011. [\[CrossRef\]](#)
32. Walia, G.S.; Carver, J.C. A systematic literature review to identify and classify software requirement errors. *Inf. Softw. Technol.* **2009**, *51*, 1087–1109. [\[CrossRef\]](#)
33. Zhang, H.; Babar, M.A.; Tell, P. Identifying relevant studies in software engineering. *Inf. Softw. Technol.* **2011**, *53*, 625–637. [\[CrossRef\]](#)
34. Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Inf. Softw. Technol.* **2015**, *64*, 1–18. [\[CrossRef\]](#)
35. Kitchenham, B.A.; Li, Z.; Burn, A.J. Validating Search Processes in Systematic Literature Reviews. In Proceedings of the Proceeding of the 1st International Workshop on Evidential Assessment of Software Technologies, Beijing, China, 8–11 June 2011; pp. 3–9.
36. Ali, N.B.; Usman, M. Reliability of search in systematic reviews: Towards a quality assessment framework for the automated-search strategy. *Inf. Softw. Technol.* **2018**, *99*, 133–147. [\[CrossRef\]](#)
37. Bailey, J.; Zhang, C.; Budgen, D.; Turner, M.; Charters, S. Search Engine Overlaps: Do they agree or disagree? In Proceedings of the Second International Workshop on Realising Evidence-Based Software Engineering (REBSE'07), Minneapolis, MN, USA, 20–26 May 2007; p. 2.
38. Chen, L.; Ali Babar, M.; Zhang, H. Towards an evidence-based understanding of electronic data sources. In Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering (EASE), Keele, UK, 12–13 April 2010.
39. Turner, M. *Digital Libraries and Search Engines for Software Engineering Research: An Overview*; Keele University: Keele, UK, 2010.

40. Wohlin, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, London, UK, 13–14 May 2014; p. 38.
41. Badampudi, D.; Wohlin, C.; Petersen, K. Experiences from using snowballing and database searches in systematic literature studies. In Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering, Nanjing, China, 27–29 April 2015; p. 17.
42. Ali, N.B.; Petersen, K. Evaluating strategies for study selection in systematic literature studies. In Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, Torino, Italy, 18–19 September 2014; p. 45.
43. Petersen, K.; Ali, N.B. Identifying strategies for study selection in systematic reviews and maps. In Proceedings of the 2011 International Symposium on Empirical Software Engineering and Measurement, Banff, AB, Canada, 22–23 September 2011; pp. 351–354.
44. Barros-Justo, J.L. ConVoc Data Extraction Form (DEF) with Raw Data. Available online: <http://doi.org/10.6084/m9.figshare.9731489.v1> (accessed on 30 July 2019).
45. Ampatzoglou, A.; Bibi, S.; Avgeriou, P.; Verbeek, M.; Chatzigeorgiou, A. Identifying, categorizing and mitigating threats to validity in software engineering secondary studies. *Inf. Softw. Technol.* **2019**, *106*, 201–230. [CrossRef]
46. Budgen, D.; Brereton, P.; Drummond, S.; Williams, N. Reporting systematic reviews: Some lessons from a tertiary study. *Inf. Softw. Technol.* **2018**, *95*, 62–74. [CrossRef]
47. Kuhrmann, M.; Fernández, D.M.; Daneva, M. On the pragmatic design of literature studies in software engineering: An experience-based guideline. *Empir. Softw. Eng.* **2017**, *22*, 2852–2891. [CrossRef]
48. Cruzes, D.S.; Dyba, T. Recommended Steps for Thematic Synthesis in Software Engineering. Available online: https://figshare.com/articles/Data_Extraction_Form_DEF_with_raw_data/9731489/1 (accessed on 30 July 2019).
49. Petersen, K.; Gencel, C. Worldviews, research methods, and their relationship to validity in empirical software engineering research. In Proceedings of the 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement, Ankara, Turkey, 23–26 October 2013; pp. 81–89.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).