

Received July 27, 2018, accepted September 7, 2018, date of publication October 5, 2018, date of current version November 8, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2874096

On the Risks and Safeguards for Requirements Engineering in Global Software Development: Systematic Literature Review and Quantitative Assessment

JOAQUÍN NICOLÁS¹, JUAN M. CARRILLO DE GEA¹, BERNABÉ NICOLÁS²,
JOSÉ L. FERNÁNDEZ-ALEMÁN¹, AND AMBROSIO TOVAL¹

¹Department of Computer Science and Systems, University of Murcia, 30100 Murcia, Spain

²TICARUM SLU, 30100 Murcia, Spain

Corresponding author: Juan M. Carrillo de Gea (jmcdg1@um.es)

This work was supported by the Spanish Ministry of Economy and Competitiveness and the European Fund for Regional Development (ERDF) through the GINSENG-UMU Project under Grant TIN2015-70259-C2-2-R.

ABSTRACT Requirements engineering (RE) is a critical process in software development which faces important risks when performed in a global software development (GSD) setting. Some of these risks are specific to GSD, while others also appear in co-localized environments but are aggravated in GSD. A systematic literature review (SLR) has been conducted to identify the risks to RE that occur in GSD, along with the safeguards with which to manage these risks. Inspired by a grounded theory approach, risks and safeguards have been elicited and grouped by means of collaborative tagging and cluster analysis techniques, and then bivariate correlation tests have been applied to measure the association between these risks and safeguards. The results of the SLR include the identification of a great variety of risks to RE in GSD (218), along with a large number of safeguards (146 in total). Starting from these results, a risks and safeguards repository that encompasses the entire state of the art on RE in GSD has been produced and is now publicly available. This repository can be used as a growing knowledge base by any organization interested in carrying out RE in GSD. The objective is to assist those organizations that are inexperienced in GSD to handle the problems that may arise when involved in a global development. The most common risks identified in the literature are related to: 1) knowledge sharing; 2) client and vendor relationships; 3) problems with process definition; and 4) communication problems. It was also found that in the literature, the risks related to management and project coordination, knowledge management and awareness, socio-cultural differences, and client-vendor distance are the ones most lacking in proposed safeguards; these are therefore concerns to which more attention should be paid.

INDEX TERMS Global software development, quantitative analysis, requirements engineering, risks and safeguards, systematic literature review.

I. INTRODUCTION

The process of globalization that affects industry has also reached Software Engineering (SE), so that increasing attention is being paid to various aspects of SE in a global environment [1]. This new software development model, in which participants are located in different countries, or even on different continents, is known as *global software development* (GSD) [2], [3]. GSD permits [4]: (1) advantage to be taken of lower labor costs in certain countries in which there are highly qualified professionals; (2) access to expert knowledge

where it is located; (3) 24-hour work days, by taking full advantage of the different time zones (the so-called *around-the-clock* or *follow-the-sun* work model); (4) global access to a shared pool of resources; and (5) an improvement as regards the proximity to certain markets.

Schmid [5] states that Requirements Engineering (RE) “is a core part of the SE lifecycle with tremendous leverage on software development success”, and claims that RE “becomes particularly difficult in the context of GSD due to the need to coordinate many different stakeholders in a

distributed setting. It is inherently more complex and difficult than RE in closely co-located settings”. RE activities are increasingly distributed across locations, organizations, participants, time and project methodologies. Such broad distribution poses new challenges and produces a severe contrast to the traditional view of RE. Challenges of communication and understanding in GSD are most pronounced during RE [6], and RE in GSD (*Global RE* hereafter) is highly problematic, leading to a large number of errors [7]. Coming back to Schmid [5], “Global RE is a challenge, and will remain a challenge for the future”.

While RE poses many unique and difficult challenges in co-located software development, the risks related to requirements are amplified in GSD; this is due to delayed visibility of open issues, as well as to uncertainties and misunderstandings [8]. From a project management perspective, Global RE risks must be specifically managed in the context of activities such as release planning [9], progress tracking [10], contract management [11], Service Level Agreement (SLA) [8], global virtual teamwork [12], and many more. Adequate RE-related risk management is thus a key success factor in GSD environments.

Cheng and Atlee’s [13] study of the state of research in RE identifies globalization as one of the *hotspots* that deserves attention in research in the RE field. Overcoming the problems that GSD implies in RE is essential if an effective globalized development of software products is to be achieved. The main risks involved in GSD concern communication problems brought about by the distribution of the development team. For example, let us consider an inherently collaborative activity such as elicitation and the *early modeling* of requirements. Tensions may occur as a result of the distance imposed by GSD, where the need arises to create a mental model of the project needs and where the requirements have to be shared by all the participants [14]. Distance should be measured in GSD not only in geographical terms but also in terms of time zones, language and culture. For Cheng and Atlee, there are two main challenges in Global RE: (1) new RE techniques need to be devised, or the current ones need to be extended, in order to support the outsourcing of development activities after RE, i.e. activities such as design, coding and testing; and (2) support techniques are needed to make it easier to carry out the requirements elicitation, modeling, negotiation, and team management in the process of Global RE; these techniques will also make it possible to manage all this activity well.

Software development and RE can be classified into two main categories according to the geographic location of work [15], [16]: (1) *onshore*, when the development is located in the same country or another not far away (the latter is also called *nearshore*); and (2) *offshore*, when the development takes place in a country which is different from that of the organization in charge of the project. Similarly, the organizational structure also has an influence. There are two main possibilities: (1) *outsourcing*, when the software developer is a third-party vendor; and (2) *insourcing*, when the

developer is an internal software IT group. Berenbach [17] studied the issue of organizational structure in more detail, identifying six different cases that range from *single site*, where there is a single site and a single command chain, to *fully distributed analysis, design and implementation*, where organizational complexity is maxed out. An important conclusion is that the problems observed tend to accumulate and grow as the degree of distribution and complexity increase.

Our study is focused on a GSD setting, which is achieved through the externalization of the development tasks to geographically remote vendors (*offshore outsourcing*), or by establishing their own software centers in remote locations, with the intent of being closer to the client’s sites (*offshore insourcing*). In both cases, the stakeholders are characterized by important diversification and geographic distribution, thus requiring appropriate collaboration, requirements change management, and coordination activities and solutions. Nevertheless, our findings are potentially applicable to other distributed software development contexts, since risks are generally not exclusive to offshore projects; they can also appear in nearshore or even onshore environments [18]. The characterization of the specific risks that can affect each organizational structure (e.g. *distributed analysis-co-located design and implementation*, *co-located analysis and design-distributed implementation*) falls outside the scope of this work.

The goal of this paper is to study the risks found in the literature in relation to Global RE, along with the safeguards that help to manage these risks. To that end, a research methodology has been designed which combines qualitative and quantitative research techniques, namely (1) a systematic literature review (SLR) [19], [20], which permits the rigorous study of the state-of-the-art in an area of knowledge; (2) grounded theory procedures [21], which help to construct an emergent, inductive model to synthesize the literature review; and (3) quantitative analysis procedures—such as collaborative tagging, clustering analysis and correlation tests—to identify relationships in that model [22].

A number of studies have been published in recent years which identify the problems that arise when developing software development activities in a global setting, and solutions to these problems have been proposed in [23]. The novelty of our research lies in two aspects:

- In an attempt to make the knowledge that can be found in the literature on Global RE more usable, the SLR has been used as a basis for the creation of a publicly available repository¹ in which all the risks and safeguards identified in the literature have been collected. This repository should be understood as a growing base of knowledge that an organization that performs GSD should tailor according to its particular know-how. We believe that this repository could be especially helpful to those organizations that are inexperienced in GSD.

¹<http://www.um.es/giisw/GSD/wiki>

- Verner *et al.* [23] state that the focus of the existing SLRs on GSD is to map the research rather than to provide evidence-based guidance. Risks and safeguards have been quantified in this paper and grouped by means of collaborative tagging and cluster analysis. Statistical, bivariate correlation tests have been applied rigorously to uncover associations between these risks and safeguards.

The structure of the remainder of this paper is as follows: Section II shows the methodology designed for this research. After that, Sections III, IV, V and VI present the *define*, *search*, *select*, and *analyze* stages of this methodology, respectively. Threats to the validity of the study also form part of Section VI. Related work is discussed in Section VII, where the contribution of this paper to the literature is stressed. Section VIII then shows our conclusions and further work. Finally, Appendix A presents the studies included in the SLR, Appendix B lists the candidate studies excluded in the SLR, and Appendix C elaborates on the descriptive analysis of the results of the SLR.

II. METHODOLOGY

The research methodology followed in this study is an adaptation of the Grounded Theory Literature Review Method proposed by Wolsfwinker *et al.* [24]. This literature review method is based on the Webster and Watson [25] framework, which has become a standard for performing literature reviews in the Information Systems (IS) field. The Grounded Theory Literature Review Method encompasses five stages (Fig. 1) which are carried out iteratively. First, the most relevant data set is identified (*Define* stage). The actual search for the studies is then performed (*Search* stage), and after that the sample of studies to be reviewed is refined (*Select* stage). In the following step, the *Analyze* stage aims to extract value from the selected studies. In the original proposal by Wolsfwinker *et al.* [24] this stage refers to the use of qualitative research methods, rooted in the Grounded Theory Method (GTM). Finally, the *Present* stage is devoted to the writing-up of the findings and insights of the research, along with the key decisions made during the review process. In a nutshell, our methodology extends this Grounded Theory Literature Review Method by using (1) SLR procedures [19] to add rigor to the definition, search and selection of studies to be reviewed (stages 1, 2 and 3 in Fig. 1); and (2) quantitative procedures to improve knowledge coding and meta-analysis during the analysis of the selected studies (stage 4).

A study by Wiesche *et al.* [26] that focuses on the use of GTM in the IS field identifies three types of research contribution: the development of a *theory*, a *rich description* of a phenomenon, or a *model* of a phenomenon. GTM is used in this paper to develop a model of a phenomenon; in our case, the phenomenon is risks that affect Global RE and the safeguards proposed for managing them. This model includes “the definitions of the relevant variables and the relationships among those variables, but it does not fully justify those relationships and specify their boundaries”. In addition, our

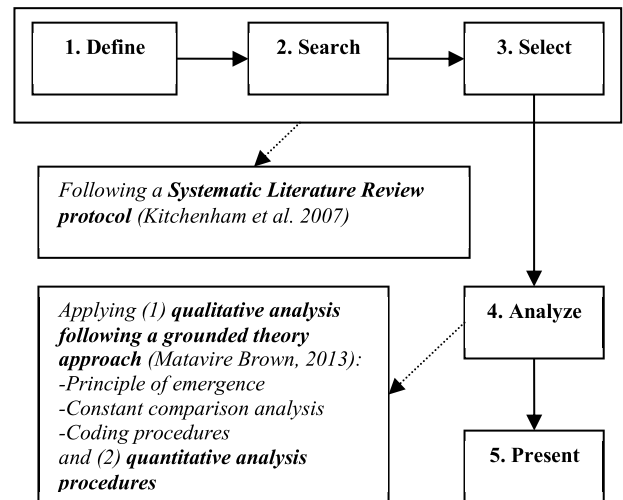


FIGURE 1. Research methodology stages, adapted from [24].

intention is to apply a set of grounded theory techniques for data analysis purposes instead of resorting to the use of a *classic* or *evolved* GTM. The former is the most common approach in the application of GTM in IS research [21].

The *Analyze* stage of our methodology follows an inductive, constructivist analysis approach that implements a set of grounded theory techniques for data analysis (Fig. 1). Following [21], these techniques are (1) the *Principle of emergence*, by avoiding the imposition of any pre-conceived structure in the model of data; (2) *Constant comparison analysis*, as data are broken down and compared to already-emerged concepts for similarities and differences; and (3) *Coding procedures*, to conceptualize and integrate data to form the model. However, our approach does not apply another GTM principle, *Theoretical sampling*, since sampling is not determined by the emerging theory but rather by the SLR protocol, which was developed prior to the *Analyze* stage. More detail on the actual implementation of the literature review methodology followed in this research is provided in the remainder of this paper, which is structured according to our five-stage methodology.

III. DEFINE STAGE

The *Define* stage consists of the identification of a suitable data set for the literature review. In our methodology this is done through the definition of the protocol followed for this SLR, which is described in Section III.A. *Concepts and definitions*, Section III.B. *Scope* and Section III.C. *Research questions*. This protocol was designed by the third author of this paper, while his supervisor (the first author) together with the second, fourth and fifth authors of this paper, reviewed the protocol, along with the included and excluded studies, and examined the results of the review with him.

A. CONCEPTS AND DEFINITIONS

Two key theoretical constructs will be used from now on; they need to be clarified beforehand: *risk* and *safeguard*. Risk is

defined as a probability that a certain event will have a negative impact on a company [27], or as a negative result whose probability of occurrence can be estimated [28], [29]. For its part, *risk management* aims to identify, assess, prioritize, and address risks, in the quest to reduce their likelihood or their effect [30], [31].

Risk management has two primary aspects [32]: (1) *risk assessment* (or *risk estimation*), which focuses on the evaluation of risk sources that could affect the outcome of the project; and (2) *risk control*, concerned with resolving the risks [33]. The latter aspect includes *risk mitigation*, which is regarded as the practices performed to reduce risk [27].

In the literature mentioned above, the term risk is conceptualized in this work as any *problem*, *issue*, *challenge* or *threat* that can adversely affect the project. The term safeguard will be understood as a risk mitigation or risk control strategy, which can be seen as a *solution*, an *aid* or a *lessening* of some risk/s. Safeguards thus include best practices, policies, guidelines, techniques, methods and tools.

B. SCOPE

In this paper, an SLR on the state-of-the-art of Global RE has been conducted to obtain:

- A description of the risks identified in the literature in relation to Global RE, understood as (1) those risks that might arise in the context of GSD which do not appear in the traditional, co-localized development model; and (2) those risks that are present in traditional settings but which worsen in the GSD realm. Risks may be caused by distance, lack of communication, organizational and cultural barriers, difficulties in project coordination, etc.
- A description of the safeguards proposed in the literature for eliminating, mitigating or managing these risks.
- The identification of gaps in the literature as regards proposed safeguards for certain risks.
- A knowledge base from which new safeguards for the risks found can be proposed.

C. RESEARCH QUESTIONS

The research questions to be answered in this SLR are the following:

- RQ1: *What risks appear in the literature in the application of Global RE?*
- RQ2: *What safeguards have been proposed in the literature to deal with the risks identified in RQ1?*

IV. SEARCH STAGE

The Search stage is undertaken to look for the papers to be used as source of information. This stage includes Section IV.A. *Sources and search strings* and Section IV.B. *Search results and protocol deviation*.

A. SOURCES AND SEARCH STRINGS

The following sources were selected to carry out this SLR:

- IEEE Digital Library (www.computer.org/portal/site/csdl/index.jsp)

- ACM Digital Library (portal.acm.org)
- ScienceDirect (www.sciencedirect.com)
- Kluwer+Springer (www.metapress.com)
- JStor (<http://www.jstor.org/>)
- EBSCOhost (<https://www.ebscohost.com/>)
- Springer-Business (<http://www.springer.com/>)
- AIS Electronic Library (<http://aisel.aisnet.org/>)
- Scopus-Social Sciences (<https://www.elsevier.com/solutions/scopus>)
- Google Scholar (scholar.google.com)

These databases index both SE and IS literature, so they are likely to provide a high level of coverage of the SLR topic. Google Scholar was selected as a source of gray literature, such as technical reports and white papers.

The keywords of the search string were derived from the research questions. Furthermore, relevant literature and various trial searches in the databases were carried out to identify alternative terms and build the following search string:

“Global” OR “offshore” OR “distributed” AND “Requirements Engineering”

Although our study focuses on RE in global, offshore scenarios, we have included the term *distributed* in the search string. In fact, these three keywords (global, offshore and distributed) have different meanings and should not be used interchangeably. As explained in Section I, offshore is a kind of global scenario; both global and offshore are more specific than distributed. There are two reasons for broadening the search string:

- Many papers in the literature refer to global and offshore scenarios simply as distributed, even though they refer to global or offshore settings. For example, Vlaar *et al.* [6] agree that distributed work “involves collaboration among teams across locational, temporal and relational boundaries to accomplish an interdependent task”, and “is characterized by geographic dispersion, reliance on electronic media, and national diversity”. Another example is provided by Schmid [5], who uses “(Globally) Distributed Requirements Engineering” to refer to “globally distributed people within the company who are involved in RE.”
- The risks encountered in distributed settings can also be found in global and offshore scenarios, where they are usually exacerbated, as discussed in Section I. It can thus be useful to find these risks in the literature to specify them in our repository.

B. SEARCH RESULTS AND PROTOCOL DEVIATION

The execution of searches for this SLR took place in June 2016. The number of studies found is presented in Table 1, grouped by source. No protocol deviation was registered in the SLR.

V. SELECT STAGE

The Select stage refines the sample of studies to be reviewed. This section includes the sub-sections V.A. *Inclusion and exclusion criteria* and V.B. *Study and data collection*.

TABLE 1. Number of studies found, candidates and those finally selected; (*) means that identical studies from different sources have been considered only once.

| Source | Studies found | Candidate studies | Selected studies |
|------------------------|---------------|-------------------|------------------|
| IEEE | 193 | 31 | 28 |
| ACM | 905 | 16 | 15 |
| MetaPress | 11 | 3 | 3 |
| Wiley InterScience | 1505 | 9 | 9 |
| JStor | 43 | 3 | 3 |
| EBSCOhost | 40 | 3 | 3 |
| Springer-Business | 330 | 2 | 1 |
| AIS | 33 | 2 | 2 |
| Scopus-Social Sciences | 149 | 17 | 16 |
| Google Scholar | 2500 | 51 | 45 |
| TOTAL | 5709 | 95(*) | 85(*) |

A. INCLUSION AND EXCLUSION CRITERIA

During the various tests concerning the search string in the different sources, it was found that in many cases simply reading the title was sufficient to decide whether or not the study should be considered as a candidate study in the SLR. As the search terms are widely used in the literature, the searches produced many studies which are not related to the scope of this SLR. When reading the title was not enough to permit a study to be included or excluded as a candidate, the abstract, the introduction and even the whole document were read if necessary.

1) INCLUSION CRITERIA

The inclusion criteria used to determine whether or not a study could be considered as a candidate study are the following:

- **i.c.1:** Studies will be accepted which deal with risks in the application of Global RE and/or propose solutions to those risks.
- **i.c.2:** Studies will be accepted if special characteristics of Global RE are described, even though these features are not specifically identified as risks, but risks can be seen as resulting from them.

2) EXCLUSION CRITERIA

The exclusion criteria used to determine which candidate studies would not eventually be selected are listed below:

- **e.c.1:** Studies not written in the English language will be excluded.
- **e.c.2:** Studies with less than four pages will be excluded.
- **e.c.3:** Studies that have been extended or updated in more recent and complete studies by the same authors will be excluded.

B. STUDY AND DATA COLLECTION

After applying the inclusion criteria, most of the studies found were not labeled as candidate studies. After subsequently

applying the exclusion criteria to the candidate studies, we obtained the studies that were eventually selected. Finally, identical studies repeated in several sources were not duplicated (see Table 1), so that 85 selected studies were eventually obtained. Selected studies are listed in Appendix A, while Appendix B shows the rejected candidate studies, together with the reason for their exclusion.

The template shown in Table 2 has been used to collect and analyze data from the selected papers, and has been adapted from that proposed by Biolchini *et al.* [34]. The form contains a section showing the objective results described by the authors, indicating the research methodology (if reported), the risks posed to Global RE, the safeguards proposed for these risks, and the problems and limitations of the study reported by the authors themselves. Finally, in a third section in Table 2, the subjective results are collected, together with the other authors' views in relation to the background of this SLR.

TABLE 2. Data collection form (adapted from Biolchini et al. [34]).

| Research ID | |
|---|---|
| Name | <i>The name of the study, in inverted commas.</i> |
| Authors | <i>Named exactly as they appear in the study.</i> |
| Journal or conference | <i>The source of the study is shown.</i> |
| Source | <i>This specifies whether the study is taken from searches of systematic reviews, from a set of available studies before the systematic review, from a related work, from a reviewed study or from a document presented in a conference which was found manually.</i> |
| Objective Results | |
| Study methodology | <i>The methodology used to arrive at the results, if reported.</i> |
| Problems found | <i>The risks (challenges, threats) identified by the authors in the application of RE to GSD.</i> |
| Proposed solutions | <i>Safeguards (solutions, best practices) to the problems found in the application of RE to GSD proposed by the authors.</i> |
| Study problems and limitations | <i>Those specifically identified by the authors.</i> |
| Subjective Results | |
| General Impressions / Possible extrapolations | <i>Our specific conclusions regarding the study.</i> |

VI. ANALYZE STAGE

The Analyze stage draws value from the selected studies of the SLR. The original Grounded Theory Literature Review Method [24] suggests the use of qualitative procedures borrowed from GTM. In contrast, the implementation of the Analyze stage in our methodology is based on both qualitative and quantitative procedures. Fig. 2 synthesizes our approach to this stage. In short, starting from the selected studies in the Select stage, first Global RE risks and safeguards are elicited

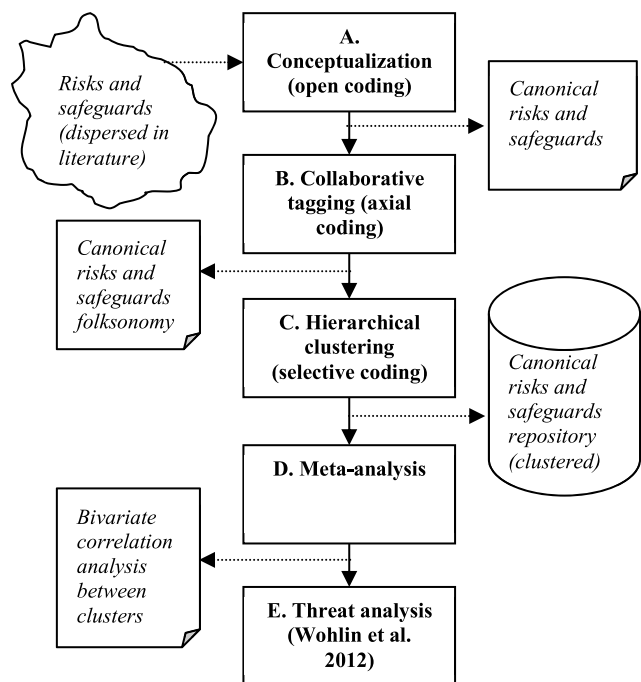


FIGURE 2. Analyze stage activities overview.

(Activity A in Fig. 2, Conceptualization). This elicitation is iteratively performed along with an emergent, bottom-up building of the so-called *canonical risks* and *canonical safeguards* from the risks and safeguards found. Once canonical risks and safeguards have been identified, our objective is to aggregate again those results. To that end, collaborative tagging allows us to build a folksonomy of the canonical risks and safeguards (Activity B); hierarchical clustering then leads to the emergent, bottom-up building of *areas of concern* to group related canonical risks (Activity C). A descriptive discussion of the concerns is also provided here. After that, bi-variate correlations are used to statistically relate canonical risks and safeguards, thus helping discover potential gaps in the research field (Activity D). Finally, threat analysis (Activity E) is carried out following the *de facto* standard approach by Wohlin *et al.* [22]. The remainder of this section provides a detailed explanation of these activities and relates them to the GTM open, axial and selective coding procedures shown in Fig. 2.

A. CONCEPTUALIZATION (OPEN CODING): KNOWLEDGE SYNTHESIS AND QUALITY CRITERIA

The analytical GTM activity of open coding is done to identify, (re-)label and/or build a set of concepts based on the excerpts drawn from the selected papers [24]. Fig. 3 presents a UML conceptual diagram designed to show the structure of the knowledge drawn from the literature. Each risk or safeguard belongs to what we call a *canonical risk* or *canonical safeguard*, respectively. These canonical risks or safeguards were defined to help us synthesize risks and safeguards found in the literature; these risks and safeguards were similar

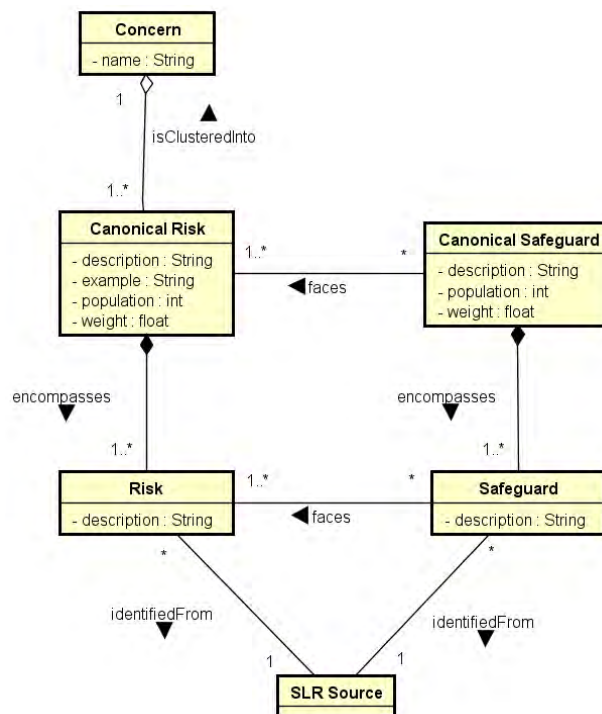


FIGURE 3. UML meta model defining the concepts used in this research.

enough to consider them within the same entity. A canonical risk or safeguard thus represents a set of risks or safeguards that are essentially the same, even if they have been found in different sources and have been written differently, or include some peculiarity despite being essentially the same. The first and third authors of this paper have identified the canonical risks and safeguards by consensus. A total of 218 risks and 146 safeguards were identified in the SLR. A total of 30 canonical risks and 29 canonical safeguards were also established.

Each canonical risk has an *example* attribute that illustrates the risk identified. Canonical safeguards have a *face* attribute in which the canonical risks will be referenced by those studies that intend to provide a solution, thereby establishing traceability between risks and safeguards. Canonical risks and safeguards also include a *population* attribute that shows how many times the associated risks or safeguards were found in the literature, and a *weight* attribute, setting an importance criterion. For this last attribute we have assigned a weight to each risk or safeguard: 0.5 to those drawn from gray literature, 1 to theoretical studies, and 2 to empirical studies. We give more value to empirical and industrial studies than to theoretical studies and gray literature, as empirical support for the majority of risks in the literature is moderate to low [23]. The weight of a canonical risk, or a canonical safeguard, will be the sum of all the weights of its related risks or safeguards. The reader can find the relationships between canonical risks and safeguards and their corresponding risks and safeguards,

respectively, in the repository². Finally, both risks and safeguards have an *identifiedFrom* attribute which identifies the source study referenced in the SLR.

B. COLLABORATIVE TAGGING (AXIAL CODING): FOLKSONOMY

The activity of axial coding is carried out to identify the interrelations between the concepts identified in open coding. This is done in this research by means of collaborative tagging. A specific series of tags was assigned to each canonical risk. Following the recommendations of Calefato et al. [35], the tags were freely chosen by the authors of this work at the time of labeling these canonical risks, and were used to describe them in a synthetic manner, according to the experience of the authors. For example:

$CanonicalRisk_1 = \{“Tag_1”, “Tag_2”, “Tag_3”\}.$

$CanonicalRisk_2 = \{“Tag_1”, “Tag_3”, “Tag_4”, “Tag_6”, “Tag_7”\}.$

Since this task was performed iteratively and collaboratively, at the end of the process a *folksonomy* was obtained, resulting in a categorization that covered all the canonical risks [36]. A folksonomy is a type of *ontology* that sits between a *thesaurus* and a *lightweight ontology* as regards complexity [37]. Fig. 4 displays a tag cloud that represents the frequency of appearance of the labels by means of the relative size of the keyword. There are no explicit hierarchies between concepts.



FIGURE 4. Tag cloud with the frequency of appearance of the labels in the folksonomy.

Formally, the set of all labels used to tag all the canonical risks is *A*:

$A = \{“Awareness”, “Collaboration”, “Communication”, “Conflict”, “Culture”, “Distance”, “Elicitation”, “Knowledge”, “Management”, “Organization”, “Process”, “Tool”, “Trust”, “Users”\}$

A semantic annotation approach was chosen to describe the canonical risks, because the flexibility of this mechanism fits well with representing the naturally overlapping concepts that can be found in the canonical risks. Furthermore, collaborative tagging allows us to produce the outcome we need for the subsequent grouping step.

C. CLUSTERING (SELECTIVE CODING): EMERGING AREAS OF CONCERN

Selective coding is an activity used to uncover, integrate and refine the categories of canonical risks. This is done in

this research through hierarchical clustering analysis. In our study, a total of 14 tags were used. The cardinality of *A* —i.e. *card(A)*— is thus 14. The set of tags assigned to the *i*th canonical risk *R_i* is a subset of *A*. For this reason, $1 \leq card(R_i) \leq 14$. In other words, each of the labels identified in the previous step of collaborative tagging may or may not be present in the categorization of a given canonical risk. These labels can therefore be encoded with a binary value 0 or 1:

$R_1 = \{“Awareness”, “Collaboration”, “Communication”\} = \{1, 1, 1, 0, \dots, 0\}.$

$R_2 = \{“Awareness”, “Communication”, “Conflict”, “Distance”, “Elicitation”\} = \{1, 0, 1, 1, 0, 1, 1, 0, \dots, 0\}.$

A binary clustering approach was thus used to group the canonical risks into concerns in accordance with their degree of affinity or relationship. In particular, the Jaccard method for binary data was chosen [38]. To interpret the outcome of a cluster analysis, a *dendrogram* is commonly used. It displays the distance at which objects (and clusters) are combined, and is considered a helpful tool in deciding on the number of clusters [39]. Looking at iFig. 5, if we cut the dendrogram at 0.6, the resulting number of clusters is six. This choice gives us a balanced and manageable number of groups.

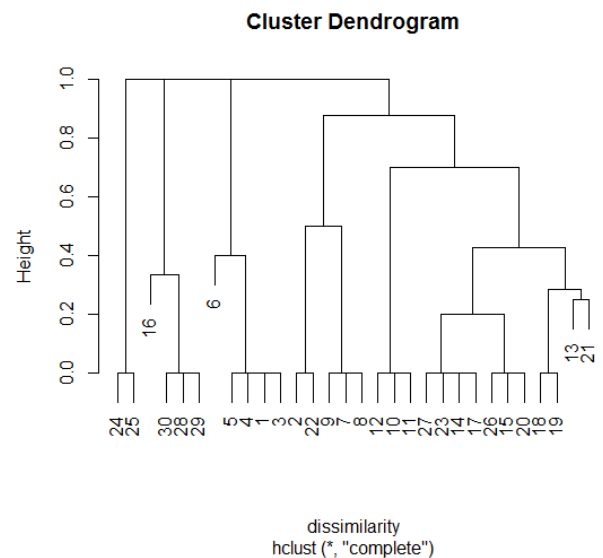


FIGURE 5. Cluster dendrogram result of binary clustering analysis.

Table 3 shows the canonical risks included in each cluster after applying the proposed method of semantic annotation and cluster analysis. The names of the clusters (*concerns* hereafter) have been assigned by the authors of this paper, inspired by the studies chosen in the SLR. The concerns are the following:

- *Communication and distance (CAD)*: caused by the stakeholders’ dispersion in space and time zones.
- *Knowledge management and awareness (KMAA)*: arising from the difficulties of managing cohesion, and the ability to share knowledge between different work groups.

²<http://www.um.es/giisw/GSD/wiki>

TABLE 3. Canonical risks by concern.

| No. | Canonical risk | Concern |
|-----|---|---------|
| 1 | Different time zones | CAD |
| 3 | Lack of informal and synchronous communication channels | |
| 4 | Lack of face to face meetings | |
| 5 | Lack of informal communication | |
| 6 | Lack of team cohesion | |
| 2 | Difficult common understanding of requirements | |
| 7 | Lack of ability to share knowledge | |
| 8 | Lack of global awareness | |
| 9 | Lack of speed when communicating changes in requirements | |
| 22 | Lack of trust | SCD |
| 10 | Socio-cultural collaborative problems | |
| 11 | Socio-cultural communication problems | |
| 12 | Regulatory differences in globally distributed settings | MAPC |
| 13 | Confidentiality and business secrets | |
| 14 | Conflict management | |
| 15 | Different terminologies and notations | |
| 17 | Distribution of work | |
| 18 | Failure to consider the costs of GSD | |
| 19 | Lack of experience in GSD | |
| 20 | Lack of coordination | |
| 21 | Lack of leadership | |
| 23 | Processes that are not well-defined | |
| 26 | Conflict between client and vendor | |
| 27 | Differences between the process models used | |
| 16 | Difficulty of extracting relevant information from stakeholders | |
| 28 | Difficulty of identifying the key users | TWSTP |
| 29 | High number of users as requirements source | |
| 30 | Lack of consideration of the expectations of end-users | |
| 24 | Lack of suitable tools for GSD | |
| 25 | Tools not employed due to lack of performance and security | |

- *Socio-cultural differences (SCD)*: resulting from interactions between groups with large social and cultural differences.
- *Management and project coordination (MAPC)*: caused by the structure of the companies involved, definition of roles and means of coordination.
- *Client-vendor distance (CVD)*: resulting from interactions among clients (e.g. *customers*) and vendors (e.g. *suppliers*) in a globalized environment.
- *Tools which support the process (TWSTP)*: caused by a lack of tools with which to support the RE process.

Two groups can be observed in Fig. 6 if we consider the number of risks that they contain. The first group corresponds to the *Management and project coordination*, *Communication and distance*, and *Knowledge management and awareness* concerns, and it contains the largest number of risks that

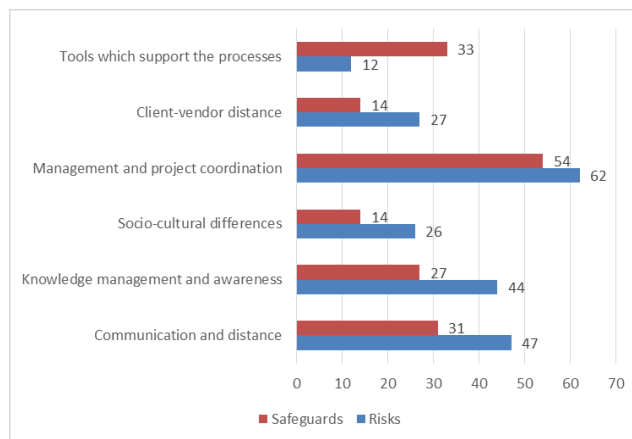


FIGURE 6. Distribution of risks and safeguards for Global RE by concern.

may lead to increased costs or the failure of a particular GSD project. The remaining concerns can be found in the second group; they consist of *Client-vendor distance*, *Socio-cultural differences* and *Tools which support the process*. We believe that the largest amount of risks is found in the first group, because risks are more likely to occur when global projects have not yet been mastered, i.e. at an early stage of maturity, while the risks in the second group are more difficult to discern and to associate with specific ones arising from GSD. In addition, *Tools which support the process* is the only concern which has more safeguards than risks associated with it, whereas safeguards for the rest of concerns are apparently far more neglected in the literature. This may be caused by the more tangible nature of the risks related to tools, in comparison to those found in other concerns, which yield a higher number of safeguards.

The numbers of risks and safeguards found in the literature are shown in Table 4, grouped by concern. In the last column we refer to the number of canonical risks with no identified safeguards in the literature. In addition, risks and safeguards and canonical risks and canonical safeguards are classified into the concerns categories, as shown in Fig. 6 and Fig. 7, respectively. Please see Appendix C for a descriptive analysis of the concerns, including their canonical risks and safeguards. Table 5 shows the canonical safeguards together with their occurrences in the literature (*Population* column).

Most of the canonical risks identified in the literature also constitute risks in co-localized environments, but they worsen in GSD because of geographical, time, linguistic and cultural distance. Furthermore, GSD canonical risks can be classified according to their *affinity* to the Global RE process; that is to say: (a) a number of the canonical risks defined in the repository are ones that concern GSD processes in general, and not only the Global RE process. But these general GSD risks also have to be taken into account in Global RE, i.e. they influence Global RE in some way; (b) there are risks that are *amplified* in RE, i.e. they are general GSD risks but they affect Global RE intensely and become especially important

TABLE 4. Risks and safeguards identified in the literature, canonical risks and safeguards, and canonical risks which are not covered in the literature studied.

| Concern | Risk | Safeguard | Canonical risk | Canonical safeguard | Not covered |
|---------|------|-----------|----------------|---------------------|-------------|
| CAD | 47 | 31 | 5 | 4 | 0 |
| KMAA | 44 | 27 | 5 | 9 | 1 |
| SCD | 26 | 14 | 3 | 7 | 0 |
| MAPC | 62 | 54 | 11 | 14 | 1 |
| CVD | 27 | 14 | 4 | 5 | 2 |
| TWSTP | 12 | 33 | 2 | 4 | 1 |

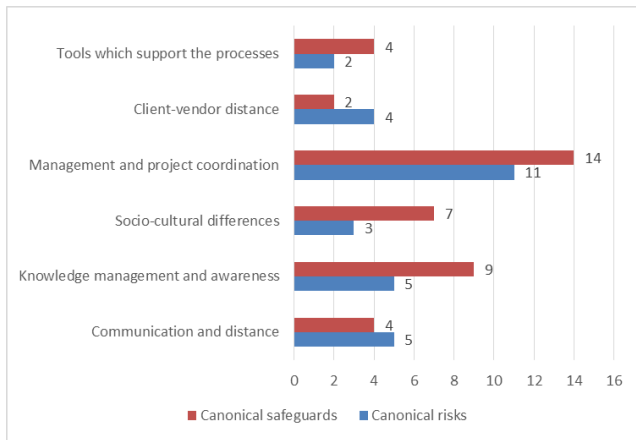


FIGURE 7. Distribution of canonical risks and safeguards for Global RE by concern.

in Global RE; and (c) there are risks that are *specific* to Global RE, i.e. they are focused on Global RE. For instance, the canonical risk *Different time zones* affects GSD processes in general, while the canonical risk *Lack of informal and synchronous communication channels* is amplified in Global RE; the canonical risk *Difficult common understanding of requirements* is specifically addressed to Global RE. Canonical risks are classified in this way in Table 6, showing the number of occurrences of each type of canonical risk in the concern (symbol \sum).

There are subtle differences between general and amplified canonical risks in some cases. The first and second authors of this paper have classified the canonical risks in Table 6 by consensus. On the other hand, canonical safeguards are classified in Table 6 according to whether they are general GSD safeguards or Global RE specific safeguards. Notice that some canonical safeguards are repeated because they are shared by more than one canonical risk.

To sum up, 9 canonical risks are general GSD risks in Table 6, but 21 canonical risks strongly affect Global RE because they are amplified in RE or specific to RE. In addition, 8 canonical risks and 8 canonical safeguards are specifically addressed to RE. As regards the canonical risks and safeguards specifically devoted to Global RE, *Knowledge management and awareness* and *Client-vendor distance* are the concerns which have the largest numbers of these.

TABLE 5. Canonical safeguards.

| No. | Canonical safeguards | Population |
|-----|--|------------|
| 1 | A common requirements notation will be chosen | 1 |
| 2 | A risk management model in the GSD | 3 |
| 3 | Bridge between two cultures | 4 |
| 4 | Client shall provide analysts direct access to members of their organization | 1 |
| 5 | Collaborative creation of requirements | 1 |
| 6 | Common processes will be suitable for the GSD model | 3 |
| 7 | Common tools will be established for every work group | 3 |
| 8 | Communication managers | 7 |
| 9 | Cultural differences of participants in the project shall be studied | 8 |
| 10 | Differences and relations between men and women will be studied | 1 |
| 11 | Domain Owner | 1 |
| 12 | Ensure there is a good team | 1 |
| 13 | Establishing of face to face meetings | 12 |
| 14 | Establish initial agreements with customers | 3 |
| 15 | Establishing of initial rules for resolving conflicts | 2 |
| 16 | Incorporate requirements attributes | 6 |
| 17 | Increase in collaboration and coordination between companies | 5 |
| 18 | It shall be established a well-defined agile leadership hierarchy | 2 |
| 19 | Knowledge sharing | 5 |
| 20 | Monitoring of development process | 1 |
| 21 | New communication channels | 9 |
| 22 | New methodologies to support GSD | 5 |
| 23 | Team backup | 1 |
| 24 | The moderator will chair requirements negotiation discussions to solve conflicts and to reach shared solutions | 4 |
| 25 | Tools with collaborative and coordination features that support distributed environments | 18 |
| 26 | Tools with communication features that support collaborative environments | 4 |
| 27 | Tools with support for iterative planning and negotiation | 2 |
| 28 | Single repository of reusable requirements | 1 |
| 29 | Use of a common language | 4 |

Moreover, the conclusion can be drawn from Table 6 that in the literature general canonical safeguards are much more numerous than canonical safeguards specific to Global RE.

D. META-ANALYSIS: CANONICAL RISKS AND SAFEGUARDS INTERPLAY

In this section, we study the results of the Analyze activities of A, B and C in a rigorous way, in an effort to gain new insights into the state-of-the-art by means of a quantitative analysis. In medicine, the majority of SLRs aim at formal meta-analysis of quantitative data [40], but this is not usual in the IS and SE fields. However, Lacity *et al.* [41] show an example of meta-analysis that aggregates both qualitative and quantitative data by coding the knowledge drawn from the literature in a set of variables, as a basis for further rigorous study of the relationships among them. In a similar fashion,

TABLE 6. Canonical risks and safeguards classified by affinity to global re.

| Concern | Canonical Risks | | | Canonical Safeguards | |
|---------|---|---|--|---|---|
| | General | Amplified in Global RE | Specific to Global RE | General | Specific to Global RE |
| CAD | -Different time zones $\Sigma 1$ | -Lack of informal and synchronous communication channels -Lack of face to face meetings -Lack of informal communication -Lack of team cohesion $\Sigma 4$ | | -New communication channels -Communication managers -Establishing of face to face meetings $\Sigma 3$ | -Collaborative creation of requirements $\Sigma 1$ |
| KMAA | | -Lack of ability to share knowledge -Lack of global awareness -Lack of trust $\Sigma 3$ | -Difficult common understanding of requirements -Lack of speed when communicating changes in requirements $\Sigma 2$ | -Establishing of face to face meetings -Monitoring of development process -Knowledge sharing -Risk management model in the GSD -Increase in collaboration and coordination between companies $\Sigma 5$ | -Incorporate requirements -Domain Owner -Single repository of reusable requirements $\Sigma 3$ |
| SCD | | -Socio-cultural collaborative problems -Socio-cultural communication problems $\Sigma 2$ | -Regulatory differences in globally distributed settings $\Sigma 1$ | -Cultural differences of participants in the project will be studied -Bridge between cultures -Team backup -Gender Analysis -Use of a common language $\Sigma 5$ | -Traceability and change analysis in the global context $\Sigma 1$ |
| MAPC | -Confidentiality and business secrets -Conflict management -Distribution of work -Failure to consider the costs of GSD -Lack of experience in GSD -Lack of coordination -Lack of leadership $\Sigma 7$ | -Processes that are not well-defined -Conflict between client and vendor -Differences between the process models used $\Sigma 3$ | -Different terminologies and notations $\Sigma 1$ | -Risk management model in the GSD -Establishing of initial rules for resolving conflicts -Increase in collaboration and coordination between companies -Tools with communication features that support collaborative environments -Common processes will be suitable for the GSD model -Common tools will be established for every work group -Ensure there is a good team -Knowledge sharing -Efficient, well-defined leadership hierarchy -New methodologies to support GSD -Tools with support for iterative planning and negotiation $\Sigma 11$ | -The moderator will chair requirements negotiation discussions to solve conflicts and to reach shared solutions -A common requirements notation will be chosen $\Sigma 2$ |
| CVD | | | -Difficulty of extracting relevant information from stakeholders -Difficulty of identifying the key users -High number of users as requirements -Lack of consideration of the expectations of end-users $\Sigma 4$ | -New methodologies to support GSD $\Sigma 1$ | -Choose an appropriate, automated prioritization requirements technique $\Sigma 1$ |
| TWSTP | -Tools not employed due to lack of performance and security $\Sigma 1$ | -Lack of suitable tools for GSD $\Sigma 1$ | | -Tools with communication features that support collaborative environments -Tools with collaborative and coordination features that support distributed environments -Tools with support iterative planning -Common tools will be established for every work group $\Sigma 4$ | |
| Total | $\Sigma 9$ | $\Sigma 13$ | $\Sigma 8$ | $\Sigma 22$ | $\Sigma 8$ |

our paper tries to aggregate data from qualitative studies in the field of Global RE; statistical correlation techniques are then used to analyze the state-of-the-practice.

An empirical evaluation of the existing relationships between canonical risks and safeguards has been conducted by means of the following calculations, where the *population* and *weight* functions are, as explained in Section VI.A, the number of times a canonical risk or safeguard was found in the literature and the modifier assigned to each canonical risk or safeguard, respectively. For each concern:

$$\text{risk}(rsk_i) = \text{population}(rsk_i) \cdot \text{weight}(rsk_i) \quad (1)$$

$$\text{safs}(rsk_i) = \sum_{j=1}^{\text{numSafs}(rsk_i)} \text{saf}_j(rsk_i) \quad (2)$$

$$\text{saf}_j(rsk_i) = \text{population}(sfg_{i,j}) \cdot \text{weight}(sfg_{i,j}) \quad (3)$$

The *risk* function is our estimation of the *relevance* of a canonical risk in the literature. It takes the form shown in (1) formally, where rsk_i refers to the i th canonical risk identified within a concern. In this calculation, we consider that the relevance of a canonical risk depends on the number of occurrences that it has in the literature. These occurrences are balanced in (1) by means of the weights described in Section VI.A, thus considering the type of the study where the risk is identified. Risks that appear in empirical studies have more weight than risks enumerated in theoretical studies, and the latter have more weight than risks described in gray literature.

The *safs* equation (2) is used to estimate the relevance of the canonical safeguards which deal with a canonical risk rsk_i . In this formula, j iterates on the existing (if so) canonical safeguards identified for rsk_i (numSafs represents this number), and saf_j returns the estimation of the relevance of a specific canonical safeguard.

Equation (3) shows how the relevance of saf_j is calculated, where $sfg_{i,j}$ refers to the j th canonical safeguard dealing with the i th canonical risk identified within a concern. We consider again that the relevance of a safeguard depends on the weighted number of occurrences it has in the literature.

Fig. 8 displays the structure of the risks and safeguards repository by means of a directed graph. The canonical safeguards and risks are represented as green and red circles, respectively. Triangles are used to depict the concerns, whereas the centralized repository is drawn with a square. All those elements make up the set of vertices of the graph. The traceability relationships between the canonical safeguards and risks, along with the concerns where they are included, are illustrated with arrows. The relative importance of the canonical risks and safeguards is correlated with the size of the circles. The outcomes of the *risk* and *saf_j* functions have been used for this purpose.

Once the operations shown in the equations above have been applied to the raw, initial data, then processed data are obtained and collected as shown in Table 7. This is a rather complex table, where we want to highlight the *risk* and *safeguards* columns first, since they show the situation regarding how canonical risks are currently dealt with by

canonical safeguards in the literature. Although it is possible to extract useful conclusions using the data alone, these columns have been divided into parts, so that two variables—one representing the risk and the other representing the associated safeguards—were obtained for each concern. The total number of variables in this study is therefore 12, namely CAD_RSK and CAD_SFG, KMAA_RSK and KMAA_SFG, SCD_RSK and SCD_SFG, MAPC_RSK and MAPC_SFG, TWSTP_RSK and TWSTP_SFG, and finally CVD_RSK and CVD_SFG. Exemplifying, CAD_RSK is the 5-tuple (66.5, 91, 110, 198, 20) and SCD_SFG is the 3-tuple (216, 6, 4).

We believe that it would be interesting to find out whether there are proper relationships between the relevance of canonical risks and the relevance of their associated safeguards in the literature. For example, there can be a canonical risk that is very relevant in the literature while its safeguards are not; this situation could flag up that more research is required to address that risk in the literature. To uncover these situations, bivariate correlation tests have been applied to measure the association between the variables. The Pearson correlation coefficient is a measure of the strength of linear dependence between two variables. It only makes sense to apply these tests to each pair of variables $\langle X \rangle_RSK$ and $\langle X \rangle_SFG$, where $\langle X \rangle$ represents a concern, since we assume that there are no other reasonable relations between the variables. The results obtained are shown as follows. To clarify their meaning, the values obtained are the *r-values*—the *r-value* indicates the strength and direction (\pm) of the correlation; the bigger the better—. The “*” or “**” signifies that the null hypothesis H_0 can be rejected (i.e. variables really correlate between them): *Communication and distance* (-0.717), *Knowledge management and awareness* (-0.317), *Socio-cultural differences* (-0.260), *Management and project coordination* (-0.079), *Client-vendor distance* (-0.355) and *Tools which support the process* (1.000**). The symbol “**” signifies that the correlation is significant at the 0.01 level (2-tailed). Some specific insights that could be drawn from these data are the following:

- The correlation between the relevance of risks and safeguards in the concern CAD - *Communication and distance* presents a negative value (-0.717), which is not high enough to be significant. In other words, most pairs of risks/safeguards might be inversely related (i.e. if a canonical risk is relevant, the canonical safeguards are not relevant, and if a canonical risk is non-relevant the canonical safeguards are relevant), but there might be deviations to this rule.
- When seeking an association between the relevance of risks and safeguards in KMAA - *Knowledge management and awareness* (-0.317), SCD - *Socio-cultural differences* (-0.260) and CVD - *Client-vendor distance* (-0.355), the anticorrelation result is not extensive. The variables do not show a clear tendency to vary together.
- There is no correlation at all between the relevance of the risks and their safeguards when considering the concern

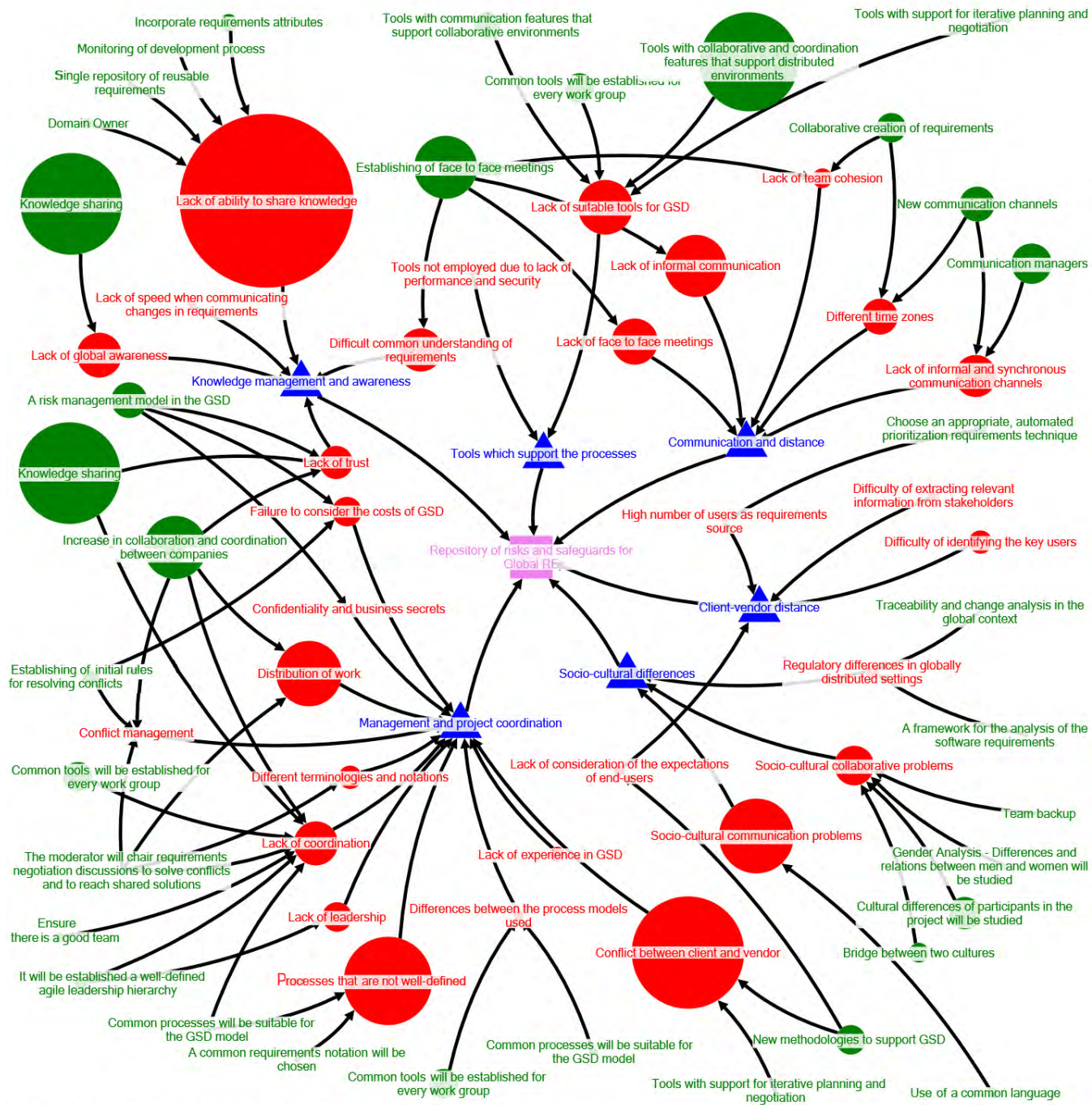


FIGURE 8. Directed graph showing the structure of the risks and safeguards repository.

MAPC - *Management and project coordination* (-0.079), thus confirming that it is necessary to study the data in greater detail. This is done through another quantitative measure, *ratio*, as discussed below—see formula (4).

- There are strong positive correlations between the variables in TWSTP - *Tools which support the process* (1.000**). In this concern, the existence of connections

between risks and safeguards were revealed, signifying that if a specific canonical risk is relevant, the canonical safeguards defined are also important; if a canonical risk is not important, the canonical safeguards are not relevant.

A formal assessment of the correlations between the previous variables is useful, but if this technique alone was applied,

TABLE 7. Numerical characterization of the canonical risks and safeguards.

| rsk_i | population(rsk_i) | weight(rsk_i) | risk(rsk_i) | numSafs(rsk_i) | \sum population($sig_{i,j}$) | \sum weight($sig_{i,j}$) | safs(rsk_i) | ratio(rsk_i) |
|---|-----------------------|-------------------|-----------------|--------------------|----------------------------------|------------------------------|-----------------|------------------|
| Communication and distance (CAD) | | | | | | | | |
| Different time zones | 7 | 9.5 | 66.5 | 2 | 12 | 18 | 216 | 3.25 |
| Lack of informal and synchronous communication channels | 7 | 13 | 91 | 2 | 14 | 21 | 294 | 3.23 |
| Lack of face to face meetings | 10 | 11 | 110 | 1 | 12 | 17 | 204 | 1.85 |
| Lack of informal communication | 11 | 18 | 198 | 1 | 12 | 17 | 204 | 1.03 |
| Lack of team cohesion | 4 | 5 | 20 | 2 | 17 | 26 | 442 | 22.10 |
| Knowledge management and awareness (KMAA) | | | | | | | | |
| Difficult common understanding of requirements | 8 | 12 | 96 | 1 | 12 | 17 | 204 | 2.13 |
| Lack of ability to share knowledge | 33 | 47 | 1551 | 4 | 7 | 12 | 84 | 0.05 |
| Lack of global awareness | 8 | 13 | 104 | 1 | 18 | 29.5 | 531 | 5.11 |
| Lack of speed when communicating changes in requirements | 2 | 2 | 4 | 0 | 0 | 0 | 0 | 0.00 |
| Lack of trust | 6 | 9.5 | 57 | 3 | 36 | 57.5 | 2070 | 36.32 |
| Socio-cultural differences (SCD) | | | | | | | | |
| Socio-cultural collaborative problems | 9 | 9 | 81 | 4 | 12 | 18 | 216 | 2.67 |
| Socio-cultural communication problems | 14 | 20.5 | 287 | 1 | 2 | 3 | 6 | 0.02 |
| Regulatory differences in globally distributed settings | 3 | 3 | 9 | 2 | 2 | 2 | 4 | 0.44 |
| Management and project coordination (MAPC) | | | | | | | | |
| Confidentiality and business secrets | 1 | 2 | 2 | 1 | 6 | 11 | 66 | 33.00 |
| Conflict management | 1 | 2 | 2 | 3 | 16 | 24 | 384 | 192.00 |
| Different terminologies and notations | 4 | 7 | 28 | 1 | 1 | 2 | 2 | 0.07 |
| Distribution of work | 11 | 20 | 220 | 2 | 13 | 19 | 247 | 1.12 |
| Failure to consider the costs of GSD | 5 | 9 | 45 | 2 | 9 | 16 | 144 | 3.20 |
| Lack of experience in GSD | 3 | 4 | 12 | 0 | 0 | 0 | 0 | 0.00 |
| Lack of coordination | 8 | 12.5 | 100 | 7 | 42 | 66.5 | 2793 | 27.93 |
| Lack of leadership | 5 | 9 | 45 | 1 | 2 | 4 | 8 | 0.18 |
| Processes that are not well-defined | 16 | 25 | 400 | 2 | 3 | 3 | 9 | 0.02 |
| Conflict between client and vendor | 18 | 36 | 648 | 2 | 6 | 10 | 60 | 0.09 |
| Differences between the process models used | 2 | 3 | 6 | 2 | 7 | 11 | 77 | 12.83 |
| Client-vendor distance (CVD) | | | | | | | | |
| Difficulty of extracting relevant information from stakeholders | 2 | 3 | 6 | 0 | 0 | 0 | 0 | 0.00 |
| Difficulty of identifying the key users | 4 | 7 | 28 | 0 | 0 | 0 | 0 | 0.00 |
| High number of users as requirements sources | 1 | 2 | 2 | 1 | 2 | 4 | 8 | 4.00 |
| Lack of consideration of the expectations of end-users | 2 | 3 | 6 | 1 | 5 | 9 | 45 | 7.50 |
| Tools which support the process (TWSTP) | | | | | | | | |
| Lack of suitable tools for GSD | 11 | 14 | 154 | 4 | 33 | 43 | 1419 | 9.21 |
| Tools not employed due to lack of performance and security | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0.00 |

the circumstances of each canonical risk by itself might go unnoticed, since correlations apply to concerns, and not to specific canonical risks. Table 7 shows a measure called *ratio*, which is calculated by using (4) to provide an individual assessment, canonical risk by canonical risk, of the amount of attention that should be paid to each one.

$$ratio(rsk_i) = safs(rsk_i) / risk(rsk_i) \tag{4}$$

If the value returned by *ratio* is greater than 1—in other words, if the *relevance* in the literature of the canonical

safeguards that deal with a canonical risk is greater than the *relevance* in the literature of that canonical risk—then it can be concluded that the *ith* canonical risk is covered by a suitable number of canonical safeguards. It seems reasonable to assume that a risk that appears a number of times in the literature should be addressed by safeguards that appear a similar number of times in the literature. In contrast, *ratio* values between 0 and 1 uncover canonical risks which should currently be demanding more attention from the scientific community to discover more safeguards to deal with them.

The value 1 corresponds to a balanced scenario, i.e. it seems that sufficient safeguards with which to tackle rsk_i can be found in the literature. In summary, an analysis of these ratios reveals the following issues:

- The ratios attained in the *Communication and distance* concern reveal that, in general, there are powerful instruments for dealing with these problems (see *ratio* column in Table 7). These values help interpret the previous result of the bivariate correlation test: the canonical risks in CAD are not as relevant (or as widely discussed in literature) as their canonical safeguards.
- In contrast, many of the canonical risks within the *Knowledge management and awareness*, *Cultural differences* and *Client-vendor distance* concerns represent a potential area of difficulties. In all these concerns, there is some imbalance between the relevance of the risks and safeguards according to the results of the correlation tests, which indicated a certain degree of inverse correlation. Firstly, in the case of *Knowledge management and awareness*, the *Lack of ability to share knowledge* (0.05) and *Lack of speed when communicating changes in requirements* (0.00) canonical risks have been identified as important barriers to GSD, but the number of convincing resources to overcome them is not relevant enough. With regard to the *Socio-cultural differences* concern, the *Socio-cultural communication problems* canonical risk (0.02) represents an appealing area for improvement and research. It is important to bear in mind that this topic is different from those found in the *Communication and distance* concern, since the former involves cultural issues. Finally, an intensive research activity is required to fill the gap between two canonical risks included in the *Client-vendor distance* concern and their safeguards. In particular, *Difficulty of extracting relevant information from stakeholders* (0.00) and *Difficulty of identifying the key users* (0.00) can be highlighted in this regard.
- The correlation analysis revealed the existence of non-related pairs where there is no correlation between variables. In *Management and project coordination*, an extensive concern with a large number of canonical risks, and indeed an insufficient number of safeguards are provided in scientific literature to deal with the canonical risks of *Different terminologies and notations* (0.07), *Lack of experience in GSD* (0.00), *Lack of leadership* (0.18), *Processes that are not well-defined* (0.02) and *Conflict between client and vendor* (0.09).
- Finally, the *Tools which support the process* concern only contains two canonical risks with the following characteristics: *Lack of suitable tools for GSD* has been thoroughly studied in the literature, especially regarding the corresponding safeguards, whereas *Tools not employed due to lack of performance and security* does not get a good result when it comes to ratio (0.00). This apparently contradicts the outcome of the correlation analysis, which showed a significant direct

correlation between variables, but actually the dimension of the latter canonical risk is minimal ($population = weight = 1$), like its safeguards which are zero.

E. THREATS TO VALIDITY

The potential threats to validity and the steps taken to mitigate or minimize them are discussed below, following the scheme put forward by Wohlin et al. [22]:

Construct Validity: Construct threats to validity in an SLR are related to the identification of primary studies [42], [43]. The principal limitation of this SLR is related to the selection of the studies included, since not all of the papers that are part of this review were empirical studies, and may therefore be lacking in validity. However, (1) the large number of studies analyzed reinforces the results and conclusions that were obtained; and (2) empirical and industrial papers have a greater weight in the repository than theoretical studies. A high quality SLR should be based on a rigorous search process. To that end, a carefully designed search string was proposed to obtain an exhaustive list of relevant primary studies. However, we are conscious that different terms exist that are closely related to the key words introduced in the search string. It is thus possible that some synonyms or homonyms relevant to the creation of terms for the search string (Section IV.A) were not considered, signifying that some additional risks or safeguards may have been omitted. This being the case, the results found may not be complete. It must also be highlighted that the references in the selected studies were not scanned to identify further studies. However, we built the search string iteratively, and we are confident that the terms used do indeed cover the set of papers related to the field of study properly. The searches were performed using IEEE Digital Library, ACM Digital Library, ScienceDirect, Kluwer+Springer, JStor, EBSCOhost, Springer-Business, AIS, Scopus-Social Sciences and Google Scholar. Nevertheless, the number of primary studies identified (85 papers) seems sufficient for us to be able to gain a deep understanding of the topic investigated. The searches made with Google Scholar provided gray literature whose quality level is not guaranteed, but a master's thesis by Ahmad and Khan [44] was eventually the only gray literature to be found and then selected in this SLR. We realize that gray literature is not as solid and reliable as scientific literature; the risks and safeguards extracted from the former therefore have a lower weight in the repository than those obtained from scientific literature.

Internal Validity: Internal validity deals with extraction and data analysis [42], [43]. Two authors performed the data extraction and classification of the primary studies, while the other authors reviewed the final results. The decision as to which data to collect and how to classify the papers was therefore performed on the judgement of the authors conducting the SLR. When conducting the SLR, articles were not excluded based on quality assessment. Even if some researchers might find it preferable to exclude those articles, we believe that including them allows us to enrich

the repository. Finally on this point, the results of the quantitative analysis are based on the repository grouping of risks and safeguards, together with concern and debatable weight assignments. All in all, we believe that our decisions are reasonable: they were agreed on by the five authors of this paper and provide a common point of view from which to analyze the state-of-the-art.

Conclusion Validity: In the case of an SLR, this threat refers to factors such as missing studies and incorrect data extraction [42], [43]. The aim is to control these factors so that an SLR can be performed by other researchers who will draw the same conclusions. Bias about selecting and classifying primary studies and analyzing data may therefore affect the interpretation of the results. To mitigate this threat, every step performed in the selection and data extraction activity was clearly described, as discussed previously. The traceability between the data extracted and the conclusions was maintained carefully. We believe that the slight differences caused by publication selection bias and misclassification should not alter the main conclusions drawn from the articles identified in this SLR. Moreover, the quantitative analysis is based on a definition of the relevance of a risk or safeguard that is debatable. A risk or safeguard that is less populated in the literature is not necessarily less important, but it is our belief that in most cases the number of appearances of a topic in the literature is a metric of the attention that the scientific community pays to it.

External Validity: External validity is related to the generalization of this study [45], [46]. The searches carried out in this SLR consider the GSD domain explicitly, and the validity of the conclusions drawn in this paper concerns the GSD context. However, the findings of this study should be validated by experts in the GSD industry. Every organization interested in the GSD realm can extend the risks and safeguards that results from this study starting from their know-how or additional information sources. In addition, some risks and safeguards drawn from the literature may be related to general issues in GSD as well as to RE. Since RE is related to many other processes, we have decided to include them all in the repository.

1) RELIABILITY ANALYSIS

of the systematic literature review We believe that this paper fulfills the checklist of features that an ideal literature review should have according to Webster and Watson [25]. Ali and Usman [45], for their part, report issues with current checklists for the quality evaluation of existing SLRs: those lists target only consistency and overlook repeatability. The more comprehensive evaluation checklist for performing and reporting SLRs in SE developed by these authors was used to perform quality assessment of our SLR. Table 8 shows the data extraction template and results.

We used automated-search alone as our search strategy. However, a set of ten databases of different types were used, including publisher and indexing databases, along with a source of gray literature (Google Scholar). Although the use

TABLE 8. Data extraction template for evaluating the conformance to quality guidelines [45].

| | |
|---|---|
| Search strategy | Automated search |
| Are keywords linked to RQs? | Yes |
| Source for synonyms | Review of relevant papers and pilot searches |
| General search string | Yes |
| Specific search string(s) | No |
| Searched fields | Not documented |
| Databases searched | IEEE Digital Library, ACM Digital Library, ScienceDirect, Kluwer+Springer, JStor, EBSCOhost, Springer-Business, AIS Electronic Library, Scopus-Social Sciences and Google Scholar |
| Rationale for the selected databases | Topic belongs to SE and IS; for this reason representative databases from both broad areas were selected |
| Search results | Yes, for each database |
| Filters on search | Till June 2016 |
| Rationale for the filters used | - |
| Known-set of papers identified? | No |
| Source for known-set of papers? | No |
| Who validated the review protocol? | The third author designed it; the first, second, fourth and fifth authors reviewed it |

of multiple search methods is often recommended, the high number of sources and their diversity provides good coverage and reduces publisher bias.

The keywords of our search string were derived from the research questions. Alternative terms and synonyms were identified by reviewing relevant, related papers and by running several trial searches. The complete general search string with logical operators was reported; the database-specific search strings and fields searched, however, were not documented during the SLR process, and cannot thus be reported. Total and database-specific search results were reported, along with the time frame filter, which is not bounded by a start date.

A known-set for validating or constructing the search string was not used in our study. The validation of the review protocol was performed by the third author (design) and the rest of the authors (review).

The previous analysis shows that the reliability of this SLR is in line with other studies of the same nature reported by Ali and Usman [45]. As acknowledged by these authors, it is not possible to repeat the search process of the majority of current SLRs, and when sufficient details are provided, there are other factors that can make the queries return different results. It is nevertheless advisable to adopt the new recommendations when carrying out this type of study if both consistency and repeatability are to be improved.

VII. RELATED WORK

The focus of this paper is rather specific; in a nutshell, the issue tackled is RE risk and safeguards in GSD. Other

papers share this narrow scope of studying a specific process in the GSD. For example, Ulziit *et al.* [46] deal with challenges and solutions for managing global software maintenance. There are, however, papers in the literature that study risk management in a more general setting, providing organizational context for the risk management process. In this regard, Aubert *et al.* [47] claim there is a need for a framework that knits multi-theory perspectives of information technology (IT) implementation, and show a reification of this framework through a multi-level theoretical model that relates risk management with higher-level decisions in the organizational economics.

The RE process in GSD is one specific aspect of a more general issue, namely IT outsourcing. Lacity *et al.* [27] take this broader perspective to perform an exhaustive review of the literature in IT outsourcing, with a focus on practice. The authors study the strategic intent behind outsourcing decisions, the risks of IT outsourcing, and how they are mitigated. This review also considers offshore IT outsourcing risks, since the authors claim that offshore IT outsourcing poses additional challenges when compared to *domestic* IT outsourcing; some of these issues are so difficult to manage that practitioners are turning to nearshore settings. The review uses a *master list* to code the similarities in the knowledge found in the literature, and to register the number of papers that deal with each topic in IT outsourcing. This approach to aggregate findings across the literature resembles the canonical risk and safeguards presented in this paper, together with their occurrences in the SLR. Lacity *et al.* [41] go a step further in the study of IT outsourcing literature, paying attention to decisions and outcomes, and pointing out research gaps in the field. The study rigorously examines evidence in the literature to identify dependent and independent variables, their frequency in the literature, and the relationships between them. One of the challenging aspects of IT outsourcing is vendor transition. Concerning this issue, Chua *et al.* [48] provide guidelines to tackle the so-called *conflicts*, by providing strategies and tactics to address them.

Coming back to RE, some researchers such as Zowghi and Damian [49], Illes-Seifert *et al.* [50], Helén [51], Kwan *et al.* [52], Ebling *et al.* [53], and Khan *et al.* [54] have already focused on the risks involved in a global development in RE. A number of reviews that offer views that complement the SLR presented in this paper can therefore be found in the literature.

Noll *et al.* [55] reviewed the GSD literature published between January 2000 and October 2009 in the *IEEEExplore* bibliographic database, in their endeavor to discover potential barriers to collaboration, as well as to identify specific practices that show how organizations can overcome these barriers. Jabangwe and Nurdiani [56] and da Silva *et al.* [57], [58] conducted SLRs on the identification of challenges in GSD along with the corresponding mitigation strategies. Jiménez *et al.* [59], for their part, conducted an SLR that focused on the initiatives related to the improvements made to the distributed software development process, and created a taxonomy of

concerns similar to those presented in our research. All of these studies look at the distributed development process in general, while our research focuses especially on RE.

Khan *et al.* [60] conducted an SLR on GSD in which the possible barriers that may be encountered by those clients who would like to outsource their software development are identified. They studied the distribution of these barriers according to various factors: (1) with regard to the size of the vendor organization (small, medium or large); (2) there may be different barriers according to the particular continent on which the client is located (Asia, North America and Europe); (3) the different study methodologies used in the literature; and finally, (4) a comparison is made between the barriers identified in the literature in the years 1990–1999 and those seen in the period of 2000–mid 2008. In a similar study, Smite *et al.* [61] conducted an SLR on GSD to describe the state-of-the-art in empirical studies of GSD, and the strength of the empirical evidence reflected in the literature. We believe that our research serves as a contribution in this regard, since we present a set of safeguards that have been identified from the relevant literature, including several techniques or best practices to be considered in GSD; these techniques and practices have additionally been classified into six different concerns.

Costa *et al.* [62] identify models and tools that support GSD, explaining that the number of studies found has increased since the year 2000, although not many tools have been developed to support it. A similar study by Prikladnicki and Audy [63] can be found, in which an SLR on process models for GSD is conducted. In our research, safeguards can be applied not only to the tools and process models surveyed by the two previous papers, but also to recommendations and best practices that help to mitigate the risks that are present in a global setting.

By means of an SLR, together with interviews of industrial experts, Nidhra *et al.* [64] identified 60 different challenges and 79 mitigation strategies for knowledge transfer in GSD environments. The authors grouped the challenges and mitigation strategies into three *factors*: (1) personnel; (2) project; and (3) technology. Sub-categories of personnel-related challenges include language barriers, cultural differences, trust, personal attributes, and staffing. Sub-categories of project-related challenges include inadequate infrastructure, problems in RE and documentation, temporal distance, changing vendor, additional costs, meeting project deadlines, coping with novelty, and communication challenges. Finally, sub-categories of technology-related challenges include challenges with tool support, and challenges with transactive memory systems.

Verner *et al.* [23] performed a tertiary study on SLRs in the GSD field. These authors found a total of 24 unique GSD SLR studies. Data extracted from each study included: (1) authors, their affiliation and publishing venue; (2) SLR quality; (3) research focus; (4) GSD risks; (5) risk mitigation strategies; and (6) for each SLR, the number of primary studies reporting each risk and risk mitigation strategy. The result of the study

is that the main GSD topics covered include (1) organizational environment; (2) project execution; (3) project planning and control; and (4) project scope and requirements. The authors elicited 85 risks and 77 risk mitigation advice items, and classified them into four categories: (1) outsourcing rationale; (2) software development; (3) human resources; and (4) project management. The largest group of risks was related to project management.

In contrast to the work already described in this section, (1) our research has a greater focus on RE; (2) a ready-to-use risks and safeguards repository comes into being as a result of this SLR; and (3) to the best of our knowledge, this is the first paper that analyzes the field quantitatively by means of statistical techniques in an effort to cluster the findings of the review, as well as to uncover some correlations between the risks and safeguards under study. López *et al.* [65] summarize an early version of the repository presented in this paper.

VIII. CONCLUSIONS AND FURTHER WORK

This paper analyzes qualitatively and quantitatively the results of an SLR concerning risks and safeguards for Global RE. Our study establishes a methodology that is rooted in GTM and quantitative analysis procedures, whose aim is to categorize knowledge in the literature concerning risks and safeguards for Global RE. It is our belief that this methodological framework, which extends the Grounded Theory Literature Review Method, is novel in Global RE; in our view it could be adapted so as to structure knowledge and develop models in other related fields of study, such as maintenance or knowledge management in GSD.

The four canonical risks that have most risks associated with them are shown in Table 9. This table shows that the most recurrent risks are: (1) knowledge sharing (e.g. *Knowledge not available in explicit form, Missing domain knowledge in offshore team*); (2) client-vendor relationships (for example, *In GSD a client may appear to be inflexible when the key issues of the project are established, In GSD vendors often assume that they do not have authority over clients, In GSD neither the client nor the vendor discuss the business model or the output of the requirements analysis*); (3) problems with process definition (for instance, *Poor control of changes; changes in the initial set of requirements, not properly managed, may result in hostile situations in relation to the final price or penalty clauses*); and finally (4) communication problems (e.g. *Language barriers, even though all the members speak English; misunderstanding can occur, as most of the language is based on cultural assumptions*).

Table 7 contains some safeguards/risk pairs with $ratio = 0$. These values highlight the risks for which no safeguards have been proposed in literature. However, not every risk has the same magnitude (see the *risk equation* in Section VI.D), and attention should thus preferably be focused on the bigger and most populated risks (e.g. the *Difficulty of identifying the key users* ratio is equal to the *Lack of speed when communicating changes in requirements* ratio (0.00), but the former risk is 7 times greater than the latter). Moreover, the

TABLE 9. The most populated canonical risks, together with their associated canonical safeguards.

| Canonical Risks | Number of times the risk was found | Canonical Safeguards |
|---------------------------------------|------------------------------------|--|
| Lack of ability to share knowledge | 33 | <i>Common processes will be suitable for the GSD model, Ensure there is a good team, Establishing of initial rules for resolving conflicts, and Increase collaboration and coordination between the two companies.</i> |
| Conflict between client and vendor | 18 | <i>New methodologies to support GSD, and Tools with support for iterative planning and negotiation.</i> |
| Processes that are not well-defined | 16 | <i>Common processes will be suitable for the GSD model, and A common requirements notation will be chosen.</i> |
| Socio-cultural communication problems | 14 | <i>Use of a common language.</i> |

Management and project coordination, Knowledge management and awareness, Socio-cultural differences and Client-vendor distance concerns are those with greatest needs in terms of safeguards to be proposed to manage canonical risks. In the *Knowledge management and awareness* concern, the canonical risk *Lack of ability to share knowledge* has a ratio close to zero (0.05), and the greatest relevance according to our study (1551). Also, in the *Management and project coordination* concern, *Conflict between client and vendor* (0.09) and *Processes that are not well-defined* (0.02) are ranked second (648) and third (400) among the canonical risks according to size, while our analysis does not show the same degree of relevance in their safeguards. The fourth canonical risk (287) is *Socio-cultural communication problems* (0.02), which is included in the *Socio-cultural differences* concern. All this suggests that these are the canonical risks to which more attention must be paid.

The results of this research are compiled in a publicly-available repository. Although our research deals with risks and safeguards collected from scientific literature, it is of a highly practical nature, which makes it useful for any company involved in GSD. For example, the repository can be useful for organizations and engineers inexperienced in GSD, who can use it as a learning tool in the GSD realm. When creating the repository, we do not intend to propose a *single* solution, in the sense of having the risks and safeguards specified and classified in the only possible way. We have attempted to create a *useful* structure instead, even though we are conscious that the results of the SLR could be grouped and specified distinctly. We should also remark that Semantic MediaWiki, a free, open-source extension to MediaWiki,³ was the technology used to implement the repository. Wikis are flexible platforms for asynchronous collaboration

³http://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki

to create content in general [66], but other knowledge representation means could have been selected for this purpose.

We believe that the repository may be a structured and extensible knowledge base that could be tailored collaboratively in accordance with the organization's best practices and know-how. We have therefore chosen a collaborative tool like Media Wiki to implement the repository. Starting from their know-how, organizations that perform GSD can introduce their own safeguards in the repository. The source attribute of these new safeguards could be *proposed*, to differentiate them from those safeguards elicited from the literature. Taking our experience as a starting point, we have introduced some *proposed* safeguards in the repository, although these have not been considered in the quantitative analysis. The current version of the repository also contains new safeguards drawn from [67], and is intended to guide organizations in the acquisition of products and services. CMMI-ACQ includes process areas such as REQM, *Requirements Management*, which helped us to extend the repository. These safeguards have not been discussed in this paper due to limitations on space, but the reader can consult them online.

APPENDIX

A. STUDIES INCLUDED IN THE SYSTEMATIC LITERATURE REVIEW

- [S1] G. N. Aranda, *et al.* (2006), "Towards a methodology for distributed requirement elicitation", presented at the XII Conference on Software Engineering and Databases, Barcelona, Spain.
- [S2] A. Ahmad and H. Khan (2008), "The importance of knowledge management practices in overcoming the global software engineering challenges in requirements understanding. Master Thesis", School of Engineering, Blekinge Institute of Technology, Ronneby, Sweden.
- [S3] J. Portillo-Rodríguez, *et al.* (2010), "Tools to Support Global Software Development Processes: A Survey", presented at the Proceedings of the 2010 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S4] D. Gumm (2007), "A model of requirements engineering at organizational interfaces: An empirical study on distributed requirements engineering", presented at the 1st International Global Requirements Engineering Workshop, Munich, Germany.
- [S5] M. Helén (2004), "Challenges in Multi-Site and Multi-Cultural Globally Distributed Software Development", Information System Science Bachelor Thesis. University of Jyväskylä, Finland.
- [S6] R. Prikladinicki, *et al.* (2005), "The role of culture in interpreting qualitative data: Methodological issues in an exploratory study of cross-cultural distributed software development", presented at the 13th Annual Cross-Cultural Meeting in Information Systems at ICIS, Las Vegas, USA.
- [S7] D. Zowghi, *et al.* (2001), "Field Studies of Requirements Engineering in a Multi-site Software Development Organization: Research in Progress", 5th Australian Workshop on Requirements Engineering, UNSW Sydney, Australia.
- [S8] D. Zowghi and D. Damian (2003), "An insight into the interplay between culture, conflict and distance in globally distributed requirements negotiations", presented at the 36th Hawaii International Conference on System Sciences, Island of Hawaii, USA.
- [S9] C. L. Iacovou and R. Nakatsu (2008), "A risk profile of offshore-outsourced development projects" *Communications of the ACM*, vol. 51 (6), pp. 89-94.
- [S10] M. Crofts, *et al.* (2004), "Global software development: The next RE frontier?", presented at the 9th Australian Workshop on Requirements Engineering, Adelaide, Australia.
- [S11] D. Damian, *et al.* (2003), "Awareness meets requirements management: awareness needs in global software development", presented at the International Workshop on Global Software Development, Portland, Oregon, USA.
- [S12] O. Uenalán, *et al.* (2008), "Using enhanced wiki-based solutions for managing requirements", presented at the 1st International Workshop on Managing Requirements Knowledge, Barcelona, Spain.
- [S13] I. Kwan, *et al.* (2007), "The effects of distance, experience, and communication structure on requirements awareness in two distributed industrial software projects", presented at the 1st International Global Requirements Engineering Workshop, Munich, Germany.
- [S14] W. J. Lloyd, *et al.* (2002), "Effectiveness of Elicitation Techniques in Distributed Requirements Engineering", IEEE Joint International Conference on Requirements Engineering.
- [S15] R. Prikladinicki, *et al.* (2003), "Requirements management in global software development: preliminary findings from a case study in a SW-CMM context", presented at the International Workshop on Global Software Development, Portland, USA.
- [S16] L. Izquierdo, *et al.* (2005), "Towards Understanding Requirements Management in a Special Case of Global Software Development: A Study of Asynchronous Communication in the Open Source Community", presented at International Workshop on Distributed Software Development (DiSD 2005), Paris, France.
- [S17] S. Jalali and C. Wohlin (2011), "Global software engineering and agile practices: a systematic review", *Journal of Software Maintenance and Evolution: Research and Practice*, vol. 24 (6), pp. 643-659.
- [S18] J. M. Bhat, *et al.* (2006), "Overcoming requirements engineering challenges: Lessons from offshore outsourcing", *IEEE Software*, vol. 23 (5), pp. 38-44.
- [S19] F. Calefato and F. Lanubile (2005), "Using the econference tool for synchronous distributed requirements workshops", presented at the 1st International Workshop on Distributed Software Development, Paris, France.
- [S20] V. Clerc (2008), "Towards architectural knowledge management practices for global software development", presented at the the 3rd International Workshop on Sharing and Reusing Architectural Knowledge, Leipzig, Germany.
- [S21] D. Damian and D. Zowghi (2002), "The impact of stakeholders? geographical distribution on managing requirements in a multi-site organization", presented at the 10th Anniversary IEEE Joint International Conference on Requirements Engineering, Essen, Germany.
- [S22] D. Damian (2007), "Stakeholders in global requirements engineering: Lessons learned from practice", *IEEE Software*, vol. 24 (2), pp. 21-27.
- [S23] D. Damian, *et al.* (2008), "On the need for mixed media in distributed requirements negotiations", *IEEE Transactions on Software Engineering*, vol. 34 (1), pp. 116-132.
- [S24] B. Decker, *et al.* (2007), "Wiki-based stakeholder participation in requirements engineering", *IEEE Software*, vol. 24 (2), pp. 28-35.
- [S25] S. Deshpande, *et al.* (2010), "Culture in global software development - A weakness or strength?", presented at the 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S26] J. Hanisch and B. Corbitt (2004), "Requirements engineering during global software development: Some impediments to the requirements engineering process - A case study", presented at the European Conference on Information Systems, Turku, Finland.
- [S27] M. Heindl and S. Biffel (2006), "Risk Management with Enhanced Tracing of Requirement Rationale in Highly Distributed Projects", presented at the 2006 International Workshop on Global Software Development for the Practitioner, Shanghai, China.
- [S28] M. Heindl, *et al.* (2007), "Requirements management infrastructures in global software development. Towards application lifecycle management with role-oriented in-time notification", presented at the International Workshop on Tool Support and Requirements Management in Distributed Projects, Munich, Germany.
- [S29] S. Islam, *et al.* (2009), "Offshore-outsourced software development risk management model", presented at the 12th International Conference on Computers and Information Technology, Dhaka, Bangladesh.
- [S30] T. Illes-Seifert, *et al.* (2007), "The challenges of distributed software engineering and requirements engineering: Results of an online survey", presented at the 1st International Global Requirements Engineering Workshop, Munich, Germany.
- [S31] F. Lanubile (2003), "A P2P toolset for distributed requirements elicitation", presented at the 1st International Workshop on Distributed Software Development, Paris, France.

- [S32] L. Paula (2010), "A Taxonomy and Visual Notation for Modeling Globally Distributed Requirements Engineering Projects", presented at the Proceedings of the 2010 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S33] E. MacGregor, *et al.* (2005), "Cultural patterns in software process mishaps: Incidents in global projects", presented at the 2005 Workshop on Human and Social Factors of Software Engineering, St. Louis, USA.
- [S34] C. Magnusson and S.-C. Chou (2010), "Risk and Compliance Management Framework for Outsourced Global Software Development", presented at the Proceedings of the 2010 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S35] V. Mikulovic and M. Heiss (2006), "How do I know what I have to do? – The role of the inquiry culture in requirements communication for distributed software development projects", presented at the 28th International Conference on Software Engineering, Shanghai, China.
- [S36] N. B. Moe and D. Smite (2008), "Understanding a lack of trust in global software teams: A multiple-case study", *Software Process: Improvement and Practice*, vol. 13 (3), pp. 217-231.
- [S37] A. Morali and R. Wieringa (2010), "Risk-based confidentiality requirements specification for outsourced IT systems", presented at the 18th IEEE International Requirements Engineering Conference, Sydney, Australia.
- [S38] S. Overhage, *et al.* (2010), "A method to evaluate the suitability of requirements specifications for offshore projects", *Business & Information Systems Engineering*, vol. 2 (3), pp. 155-164.
- [S39] P. Liang, *et al.* (2009), "Requirements Reasoning for Distributed Requirements Analysis Using Semantic Wiki", presented at the Proceedings of the 2009 Fourth IEEE International Conference on Global Software Engineering, Limerick, Ireland.
- [S40] F. Salger, *et al.* (2010), "Knowledge Transfer in Global Software Development - Leveraging Ontologies, Tools and Assessments", presented at the Proceedings of the 2010 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S41] V. Sinha, *et al.* (2006), "Enabling Collaboration in Distributed Requirements Management", *IEEE Software*, vol. 23 (5), pp. 52-61.
- [S42] D. Smite (2006), "Requirements management in distributed projects", *Journal of Universal Knowledge Management*, vol. 1 (2), pp. 69-76.
- [S43] C. Solis and N. Ali (2010), "Distributed Requirements Elicitation Using a Spatial Hypertext Wiki", presented at the Proceedings of the 2010 5th IEEE International Conference on Global Software Engineering, Princeton, USA.
- [S44] S. Lohmann, *et al.* (2008), "Semantifying requirements engineering – The SoftWiki approach", presented at the I-SEMANTICS '08, Graz, Austria, 2008.
- [S45] D. Zowghi and D. Damian (2003), "RE Challenges in Multi-site Software Development Organisations", *Requirements Engineering*, vol. 8 (3), pp. 149-160.
- [S46] G. N. Aranda, *et al.* (2005), "Towards a cognitive-based approach to distributed requirement elicitation process", presented at the VIII Workshop on Requirements Engineering, Porto, Portugal.
- [S47] L. T. Lopes, *et al.* (2005), "Requirements Specification in Distributed Software Development – A Process Proposal", in 38th Hawaii Intl. Conf. on System Sciences (HICSS'05), Hawaii, USA.
- [S48] B. Berenbach (2006), "Impact of organizational structure on distributed requirements engineering processes: Lessons learned", presented at the 1st International Workshop on Global Software Development for the Practitioner, Shanghai, China.
- [S49] C. Scharff (2011), "An evolving collaborative model of working in students' global software development projects", presented at the Proceedings of the 2011 Community Building Workshop on Collaborative Teaching of Globally Distributed Software Development, Waikiki, Honolulu, USA.
- [S50] M. Daneva, *et al.* (2013), "Agile requirements prioritization in large-scale outsourced system projects: An empirical study", *Journal of Systems and Software*, vol. 86 (5), pp. 1333-1353.
- [S51] F. Calefato, *et al.* (2012), "Computer-mediated communication to support distributed requirements elicitation and negotiations tasks", *Empirical Software Engineering*, vol. 17 (6), pp. 640-674.
- [S52] S. I. Hashmi, *et al.* (2013), "A communication process for global requirements engineering", presented at the Proceedings of the 2013 International Conference on Software and System Process, San Francisco, USA.
- [S53] L. Han, *et al.* (2012), "A Survey of RE-specific Wikis for Distributed Requirements Engineering", presented at 2012 Eighth International Conference on Semantics, Knowledge and Grids (SKG), Beijing, China.
- [S54] K. Stapel and K. Schneider (2012), "Managing knowledge on communication and information flow in global software projects", *Expert Systems*, vol. 31(3), pp. 234-252.
- [S55] S. Betz, *et al.* (2012), "Knowledge transfer in offshore outsourcing software development projects: an analysis of the challenges and solutions from German clients", *Expert Systems*, vol. 31(3), pp. 282-297.
- [S56] M. Niazi, *et al.* (2012), "An Empirical Study Identifying High Perceived Value Requirements Engineering Practices in Global Software Development Projects", presented at the Seventh International Conference on Software Engineering Advances (ICSEA 2012), Lisbon, Portugal.
- [S57] T. Fancott, *et al.* (2012), "Towards Next Generation Requirements Engineering", presented at International Conference on Social Informatics (SocialInformatics), Alexandria, USA.
- [S58] J.L. Fernández-Alemán, *et al.* (2016), "Effects of Using Requirements Catalogs on Effectiveness and Productivity of Requirements Specification in a Software Project Management Course", *IEEE Transactions on Education*, vol. 59 (2), pp. 105-118.
- [S59] K. Dullemond, *et al.* (2014), "Collaboration Spaces for Virtual Software Teams", *IEEE Software*, vol. 31 (6), pp. 47-53.
- [S60] J.M. Carrillo de Gea, *et al.* (2016), "Co-located and distributed natural-language requirements specification: traditional versus reuse-based techniques", *Journal of Software: Evolution and Process*, vol. 28 (3), pp. 205-227.
- [S61] S. Ghobadi, L. Mathiassen (2016), "Perceived barriers to effective knowledge sharing in agile software teams", *Information Systems Journal*, vol. 26 (2), pp. 95-125.
- [S62] P. Parviainen, M. Tihinen (2014), "Knowledge-related challenges and solutions in GSD", *Expert Systems*, vol. 31 (3), pp. 253-266.
- [S63] I. Inayat, *et al.* (2015), "A systematic literature review on agile requirements engineering practices and challenges", *Computers in Human Behavior*, vol. 51, Part B, pp. 915-929.
- [S64] P. Holtkamp, *et al.* (2015), "Soft competency requirements in requirements engineering, software design, implementation, and testing", *Journal of Systems and Software*, vol. 101, pp. 136-146.
- [S65] P. Laurent, *et al.* (2013), "Evaluating the Effectiveness of a Collaborative Requirements Engineering Modeling Notation for Planning Globally Distributed Projects", presented at the International Conference on Software Engineering Research and Practice (SERP), Athens, Greece.
- [S66] V. N. Viithana (2015), "Scrum Requirements Engineering Practices and Challenges in Offshore Software Development", *International Journal of Computer Applications*, vol. 116(22), pp. 43-49.
- [S67] H. Hiisilä, *et al.* (2015), "Challenges of the Customer Organization's Requirements Engineering Process in the Outsourced Environment – A Case Study", presented at the 21st International Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2015), Essen, Germany.
- [S68] M. Spichkova, H. Schmidt (2015), "Requirements Engineering Aspects of a Geographically Distributed Architecture", presented at 10th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE), Barcelona, Spain.
- [S69] J. Iqbal, *et al.* (2013), "A Framework to Improve the Requirements Engineering Process for Software Development Outsourcing", presented at 22nd Australian Software Engineering Conference, Melbourne, Australia.
- [S70] C. Sillaber, R. Breu (2014), "The Impact of Knowledge Sharing Platforms in Distributed Requirements Engineering Scenarios: A Systematic Review", presented at the 8th International Conference on Knowledge Management in Organizations: Social and Big Data Computing for Knowledge Management, Springer, Dordrecht.
- [S71] Paul W. L. Vlaar, *et al.* (2008), "Cocreating understanding and value in distributed Work: How Members of Onsite and Offshore Vendor Teams Give, Make, Demand, and Break Sense", *MIS Quarterly*, vol. 32 (2), pp. 227-255.
- [S72] H. Bruun, S. Sierla (2008), "Distributed Problem Solving in Software Development: The Case of an Automation Project", *Social Studies of Science*, vol. 38 (1), pp. 133-158.
- [S73] J. Dibbern, *et al.* (2008), "Explaining Variations in Client Extra Costs between Software Projects Offshored to India", *MIS Quarterly*, vol. 32 (2), pp. 333-366.
- [S74] M. Wasim Bhatti, A. Ahsan (2015), "Global software development: an exploratory study of challenges of globalization, HRM practices and process improvement", *Review of Managerial Science*, vol. 10(4), pp. 649-682.

- [S75] J. Gasparas, E. Monteiro (2009), “Counteracting forces in implementation of IS-enabled global business processes”, presented at the 17th European Conference on Information Systems, ECIS. Verona, Italy.
- [S76] T.D. Breaux, et al. (2009), “A distributed requirements management framework for legal compliance and accountability”, *Computers & Security*, vol. 28(1-2), pp. 8-17.
- [S77] I. Inayat, S. Salwah Salim (2015) “A framework to study requirements-driven collaboration among agile teams: Findings from two case studies”, *Computers in Human Behavior*, vol. 51, Part B, pp. 1367-1379.
- [S78] M. Abejide Adegoke, et al. (2015), “An empirical study of automated classification tools for informal requirements in large scale systems”, *International Journal of Business Information Systems*, vol. 19(3), pp. 324-341.
- [S79] S. Scheibmayr, et al. (2009), “An Experimental, Tool-Based Evaluation of Requirements Prioritization Techniques in Distributed Settings”, presented at the 15th Americas Conference on Information Systems (AMCIS 2009), San Francisco, USA.
- [S80] N. Dhruv, et al. (2008), “Project Quality of Offshore Virtual Teams Engaged in Software Requirements Analysis: An Exploratory Comparative Study”, *Journal of Global Information Management*, vol. 16 (4), pp. 24-45.
- [S81] V. Vijayakumar, et al. (2008), “Storytelling – a Method to Start Knowledge Transfer in Offshore Software Development Teams”, presented at the 9th European Conference on Knowledge Management (ECKM), Southampton, UK.
- [S82] M. Spichkova, et al. (2015), “Structuring diverse regulatory requirements for global product development”, presented at the IEEE 8th International Workshop on Requirements Engineering and Law (RELAW), Ottawa, Canada.
- [S83] M. Nordio, et al. (2009), “The Role of Contracts in Distributed Development”, presented at Software Engineering Approaches for Offshore and Outsourced Development: Third International Conference (SEAFOOD 2009), Zurich, Switzerland.
- [S84] T.O.A. Lehtinen, et al. (2015), “Why the Development Outcome Does Not Meet the Product Owners’ Expectations?”, presented at Agile Processes in Software Engineering and Extreme Programming: 16th International Conference, XP 2015, Helsinki, Finland.
- [S85] K. Schmid (2014), “Challenges and solutions in global requirements engineering - A literature survey”, presented at Software Quality. Model-Based Approaches for Advanced Software and Systems Engineering: 6th International Conference, SWQD 2014, Vienna, Austria.

B. EXCLUDED CANDIDATE STUDIES

Table 10 shows the candidate studies that were excluded from the SLR.

C. DESCRIPTIVE ANALYSIS OF THE CONCERNS

The following sub-sections present the risks and safeguards in Global RE that have been found in the SLR. Grouped into the six concerns of our research, we first present the canonical risks (thus answering RQ1), followed by the corresponding safeguards (answering RQ2).

1) COMMUNICATION AND DISTANCE

Communication in Global RE is a key factor in the project’s success. The following are the canonical risks that result from our analysis of the literature and that have to do with increased distance between the stakeholders leading to a worsening in communication. Each canonical risk is described briefly below, together with its canonical safeguards. A complete description of the risks and safeguards of the concern can be found in the repository.

Different time zones [16], [17], [44], [68]–[72], a factor that causes delays in urgent matters and that is due to time-area differences between different groups [44]. Sometimes, the limited time window available for exchanging knowledge

in GSD means that synchronous virtual meetings are held between sites with large differences in timetables, so the outcome is often inconvenient or confusing; the meeting will be too late in some places, or too early in others [73]. The canonical safeguards which we have found are the following: (1) *Collaborative creation of requirements* [74], [75] and (2) *New communication channels* [16], [44], [71], [76]–[78]. *Collaborative creation of requirements* implies that requirements documents will be composed collaboratively by remote analysts and team heads. *New communication channels*, for its part, is a safeguard that includes the use of asynchronous communication such as e-mail [44] to establish communication mechanisms so that important questions are answered without fail [70], or synchronous communication such as regular meetings via video conference [16] during any available work period in which groups overlap. In this regard, Damian [16] finds that small distributed teams have seen weekly meetings supported by video- or teleconferencing to be successful for requirements gathering, validation, and management activities. An attempt to create open lines of communication between user roles is made by Clerc [70], who proposes the creation of a clear organizational structure with people in charge of communication.

Lack of informal and synchronous communication channels [7], [16], [49], [75], [79], which is a risk wherever poorly-defined/poorly-used communication channels may appear in GSD. The canonical safeguards here are (1) *Communication managers* [70], [71], [80], [81] and again (2) *New communication channels* [16], [44], [71], [76]–[78]. *Communication managers* imply the modeling of the flow of information (including documents or verbal communication) of a distributed software project [81]. It is also recommended that a repository of artifacts, defining a storehouse of important decisions, as well as the reasons for such decisions [70], should be established. Urgent response is also needed to set up communication mechanisms so that the very important and urgent questions can be answered quickly [70]. Another way is to conduct reverse presentations, i.e. let the vendor or client explain in their own words if they have understood what you have just explained [71].

Lack of face to face meetings [15], [44], [50], [51], [68], [71], [77], [82], [83], which causes misunderstandings, lack of trust and repetition of work. The canonical safeguard that deals with this risk is *Establishing of face to face meetings* [44], [50], [70], [71], [82], [84], which includes proposals for managing this canonical risk such as having an initial meeting [70], [84], increasing the number of face to face interactions [50], [70], [82], and establishing a local office near the partner company [70]. Weekly meetings to share knowledge are advisable, with the following pattern [50]: “At least NUMBER_OF_MEETINGS meetings will be carried out in each requirements engineering process iteration, in which coordinator and team heads will take part”, where NUMBER_OF_MEETINGS should be replaced with the appropriate value (i.e. a natural number).

TABLE 10. Candidate studies excluded from the systematic literature review.

| Source | Study | Exclusion criteria (Section V.A.2) |
|--------------------------------------|--|------------------------------------|
| Google Scholar | Didar Zowghi, "Does Global Software Development Need a Different Requirements Engineering Process?", in: Proceedings of International Workshop on Global Software Development – ICSE 2002, Orlando, Florida, USA, 2002, 53-55. | e.c.2 |
| IEEE Digital Library | Steffen Lohmann, Philipp Heim, Kim Lauenroth, "Web-based Stakeholder Participation in Distributed Requirements Elicitation", in: 16 th IEEE International Requirements Engineering Conference 2008 | e.c.2 |
| Google Scholar | Daniela Damian, "The Study of Requirements Engineering in Global Software Development: as Challenging as Important", in: Intl. Conf. on Software Engineering (ICSE'02), Orlando, USA. | e.c.3 |
| IEEE Digital Library, Google Scholar | Leandro Lopes, Rafael Prikladnicki, Jorge Audy, "Distributed Requirement Specification: Minimizing the Effect of Geographic Dispersion", in: ICSE 2002 | e.c.3 |
| IEEE Digital Library, Google Scholar | Daniela Damian, "An empirical study of requirements engineering in distributed software projects: is distances negotiation more effective?", in: Asian Pacific Software Engineering Conference 2001 | e.c.3 |
| Google Scholar | Daniela Damian, "The Use of a Multimedia Group Support System for Distributed Software Requirements Meetings", in: HICCS'35 2002 | e.c.3 |
| Google Scholar | Daniela Damian, "A Research Methodology in the Study of Requirements Negotiations in Geographical Distributed Software Teams", in: 11th IEEE International Requirements Engineering Conference 2003 | e.c.3 |
| ACM Digital Library | Ahmed D. Alharthi, Maria Spichkova, Margaret Hamilton, "Requirements Engineering Aspects of eLearning Systems", in: ASWEC '15 Vol. II September 28 - October 01, 2015, Adelaide, SA, Australia. | e.c.2 |
| Springer-Business | Matthias Jarke, Kalle Lyytinen, "High Impact Requirements Engineering", in: Business & Information Systems Engineering, Vol. 2: Iss. 3, 123-124, 2010. | e.c.2 |
| Scopus-Social Sciences | Mariele Hagen, Berit Jungmann, Kim Lauenroth, "Wiki-gestütztes verteiltes Requirements Engineering für große Stakeholdergruppen", in GeNeMe '07, Dresden, Germany. | e.c.1 |

Lack of informal communication [6], [15], [51], [68], [70], [71], which affects the negotiation of requirements, the relationships among stakeholders, and the bonds of trust between stakeholders. The canonical safeguard that deals with this risk is *Establishing of face to face meetings* [44], [50], [70], [71], [82], [84]. New forms of communication, such as the use of video conferences [77], together with an increase in informal communications by e-mail, chat or telephone [44], are suggested.

Lack of team cohesion [17], [49], [51], [81]. Distance can cause timidity and passivity, even in skilled experts [17], loss of team cohesion, and a lack of trust [51]; may even result in certain members of the distributed team in GSD vying for power and influence [49]. Collaborative work and face to face meetings are key to enhancing trust and a sense of being a team [70], [82], [84]; in this respect, the canonical safeguards that we have found are (1) *Collaborative creation of requirements* [74], [75] and (2) *Establishing of face to face meetings* [44], [50], [70], [71], [82], [84].

2) KNOWLEDGE MANAGEMENT AND AWARENESS

Appropriate knowledge management and clear awareness of the workspace are key objectives within any software project. While these are difficult targets to achieve in a traditional environment [85], a number of additional risks may arise in a GSD environment. These risks are included in the following canonical risks:

Difficult common understanding of requirements [6], [17], [74], [80], [86]–[89], as it is difficult to achieve a common comprehension of requirements in GSD. Developers often realize that the rationale of the requirements is not clear in GSD, and they do not know exactly why they were formulated as they were. It is thus quite common for developers to misunderstand the requirements in GSD, thereby causing delays and needing additional effort to explain them [87]. Analysts who are used to non-distributed developments may have problems in creating specifications that are detailed enough for outsourcing [17]. The canonical safeguard that corresponds to this risk is *Establishing of face to face meetings* [44], [50], [70], [71], [82], [84], which enables face-to-face cooperation and communication, as well as social and cultural workshops. It is advisable to start projects with a kind of *boot camp* [71], i.e. to establish initial relationships by means of a face to face meeting whose goal is for those attending to obtain a shared understanding of the project context and the goals to be achieved [70]. Ahmad and Khan [44] propose a visual representation of the requirements, together with new means of communication, such as *intranets* or discussion forums.

Lack of ability to share knowledge [5], [7], [14], [16], [17], [50], [71], [72], [75], [78], [90]–[93]. Awareness in co-located projects benefits from informal communication flows that sometimes transmit the changes of the state of the project quicker than the formal documentation does. However, the projects in GSD do not receive the benefits

of the social mechanisms and natural processes that traditional developments possess [14]. In GSD the lack of trust and inability to share knowledge makes it more difficult to reach a methodology that fosters the exchange of critical information during requirements management [16]. In addition, analysts and developers in GSD might know nothing about the application domain, especially in their first collaborations [17]. Knowledge is sometimes not available in explicit form [71], and the offshore team may have a feeling that knowledge about the domain is missing [71]. Four canonical safeguards have been identified in the literature to deal with this risk: (1) *Incorporate requirements attributes* [87], [94], [95]; (2) *Domain Owner* [94], [96]; (3) *Single repository of reusable requirements* [90]; and (4) *Monitoring of development process* [7]. Concerning the first safeguard, Overhage et al. [95] propose that those colleagues with a greater domain knowledge should interpret requirements specifications in a more accurate manner, setting the necessary attributes to the requirements. Regarding the second safeguard, Daneva et al. [94] propose the creation of a new role called *Domain Owner*, an individual who is responsible for accumulating knowledge concerning the client's business domain, sharing it among team members, and packaging it in a form that is reusable in future projects. This new role raises the costs of the project but improves the transfer of knowledge. The third safeguard, *Single repository of reusable requirements*, means that an improvement can be made to the specification process by using learning based on reusable requirements catalogs [90]. Finally, the fourth safeguard, *Monitoring of development process*, is related to status reporting practices [7], i.e. reporting formats, reporting channels, decision authorities, and problem-solving practices should be defined (distributing drafts of schedules and task assignments for each incremental release, creating weekly task reports and meetings within a subgroup, along with delivery reports).

Lack of global awareness [5], [14], [16], [71], [95], [97] refers to the risk that people do not have a shared, unified view of the project. For example, people often work on the same requirements in GSD without having a clear idea of who is working on what. Another example is the need to manage and document the tacit client support knowledge that accumulates gradually throughout various systems [98]. This is a rather complex canonical risk, and a number of the collected safeguards could be considered to help indirectly in this regard. We have specified one canonical safeguard: *Knowledge sharing* [44], [71], which is intended to leverage the interaction of the team with other groups [44]. Global awareness can be promoted by implementing a knowledge repository, as part of the business process and existing tools in use. The use of knowledge maps and the installation of a knowledge management team can also be useful [71].

Lack of speed when communicating changes in requirements [14], [99]. This refers to the delay in communicating the changes in the requirements to the stakeholders. Changes in requirements in GSD are usually not reported to relevant

stakeholders as soon as they occur [99]. No specific safeguard with which to deal with this risk has been found in the literature.

Lack of trust [44], [95], [100]–[102], which refers to the lack of trust between groups in a global development environment, owing either to the offshore company's lack of expertise or to its lack of a business model. A competitive rather than a cooperative relationship may be established in GSD, and this can cause a lack of sense of being a team and bring about a drop in motivation. Several safeguards to increase trust within and between the groups have been found in the literature: (1) *Risk management model in the GSD* [7], [71], [103]–[106]; (2) *Increase in collaboration and coordination between companies* [5], [7], [44], [52], [95]; and again (3) *Knowledge sharing* [6], [7], [44], [71], [88], [102], [107], [108].

3) SOCIO-CULTURAL DIFFERENCES

In a traditional, co-localized development it is common for all the members in a working group to share the same or similar cultures, or to have a common understanding of the cultural knowledge that may be considered as *preponderant*. In a global environment, on the other hand, a variety of cultures with different degrees of compatibility between them may exist. For instance, cultural differences in terms of power distance, information system designer values, and an active versus passive working attitude critically influence offshore outsourcing success [109]. Cultural distance is also a factor that influences outsourcing decisions [110] and the allocation of work packages in GSD [111]. There are *soft internationalization competencies* such as cultural awareness, the understanding of other people's perspectives, needs and values, and foreign language skills, all of which play an important role in Global RE [112]. The canonical risks that we have identified in the literature are the following:

Socio-cultural collaborative problems [50], [51], [71], [112]–[114] may arise between different groups as a result of differences in education, culture or experience in GSD projects. The following canonical safeguards have been identified in the literature: (1) *Cultural differences of participants in the project will be studied* [7], [71], [115], [116], so that conducting a preliminary study on the degree of compatibility between cultures could prevent future problems that come about as a result of cultural differences; (2) *Bridge between two cultures* [5], [71], [116], [117], because the project can be aided by a person who has lived in both cultures for a long enough period of time; (3) *Team backup* [116], i.e. creating *backup teams* in different locations can assist in dealing with unforeseen events that are produced by cultural and festive events, thus helping to provide a 24/7 support throughout the year; and finally (4) *Gender Analysis - Differences and relations between men and women will be studied* [116], since identifying the role of female team members within working teams in other cultures can help to mitigate potential collaboration risks between the working groups.

Socio-cultural communication problems [5], [15], [44], [51], [70], [82], [83], [95], [100], [108], [112], [118]. This canonical risk can be caused by cultural differences (since each culture has a different way of thinking and tackling problems), as well as by linguistic and vocabulary differences that may result in incoherent requirements. As one way of reducing this risk, the canonical safeguard *Use of a common language*, asks for the formal establishment of a common language between the working groups [50], [95].

Regulatory differences in globally distributed settings [5], [112], [119], [120], due to the diversity of legal requirements in different contexts (countries, organizations, situations, etc.). Developing a system for different contexts implies recognizing that the regulatory, legal requirements for the system can differ from one location to another. This diversity should be managed in a systematic way [120], taking into account the variability in compliance, and avoiding contradictions. In this sense, Spichkova et al. [120] propose the following safeguard: *Traceability and change analysis in the global context*. In a nutshell, this safeguard means the use of a framework for requirements structuring, as well as traceability and change analysis in Global RE.

4) MANAGEMENT AND PROJECT COORDINATION

A number of canonical risks that may arise in management and project coordination in GSD have been identified in the literature. This is a troublesome issue, since it is not straightforward to initiate changes in a network of loosely-coupled operators with different cultures, management structures, compensation and motivation models [98]. The canonical risks encountered in the literature are listed below. In the particular cases of some of these risks, safeguards to manage them properly in GSD have not been found.

Confidentiality and business secrets [105]. Confidentiality and protection of trade secrets become potential risks in outsourcing, since vendor companies do not usually wish to reveal their infrastructure and methodology to their clients and thereby allow them to make an assessment of potential confidentiality risks. To deal with this risk, the following canonical safeguard has been identified: *Risk management model in the GSD* [7], [71], [103]–[106]. In this regard, Morali and Wieringa [105] introduce the CRAC (*Confidentiality Risk Assessment and Comparison*) method, which has been adapted to distributed environments and is called CRAC++. This method gives explicit consideration to confidentiality risks and protection of business secrets.

Conflict management (Damian & Zowghi, 2002) highlights difficulties in managing conflicts between groups or companies in GSD, as well as in maintaining open discussions. Three canonical safeguards have been found in the literature: (1) *Establishing of initial rules for resolving conflicts* [82], [89], [95]; (2) *Increase in collaboration and coordination between companies* [5], [7], [44], [52], [95], [102]; and (3) *The moderator will chair requirements*

negotiation discussions to solve conflicts and to reach shared solutions [116].

Different terminologies and notations [6], [17], [50], [95]. It is possible for remote work groups in GSD to use different requirements terminology or notations [50], thus leading to requirements or even contracts which are misunderstood or interpreted in different ways [95]. Vlaar et al. [6] also address the so-called *experience asymmetries*: when offshore team members do not have sufficient knowledge of particular technologies and programming languages, they do not fully understand the terminologies and technological jargon used in the requirements documents. Experienced members offshore generally find it easier to understand requirements, and do not need as much communication compared to novices. A canonical safeguard has been identified in the literature: *The moderator will chair requirements negotiation discussions to solve conflicts and to reach shared solutions* [116]. When the onsite team reports and updates offshore team members on the discussions they have had with clients, onsite and offshore team members can jointly develop better solutions for the issues that were brought up by clients. They start to rely more on each other's experience and knowledge, viewing problems and potential solutions from multiple perspectives, and co-creating new understandings [6].

Distribution of work [79], [101], [121] is related to the greater complexity involved when creating a balanced distribution of work among all the groups involved in a GSD project. Two canonical safeguards have been identified: (1) *Increase in collaboration and coordination between companies* [5], [7], [44], [52], [95], [102], so that the workload is more equitable; and (2) *The moderator will chair requirements negotiation discussions to solve conflicts and to reach shared solutions* [116], i.e. the use of a moderator who mediates between the various companies in the effort to solve conflicts in the distribution of workload among the groups.

Failure to consider the costs of GSD [71], [75], [100], [107]: a lack of estimates of extra cost may arise in a global environment; this goes beyond the contract-based payments to the vendor for traveling, accommodation, duplication of the project structure, controlling vendor performance, coordinating project resources, transferring knowledge to vendor personnel, and specifying software requirements. The client can usually incur post contractual extra costs for four types of activities: requirements specification and design, knowledge transfer, control, and coordination. These extra costs are quite often underestimated [107]. Three canonical safeguards to solve this risk have been identified in the literature: (1) *Risk management model in the GSD* [7], [71], [103]–[106], encompassing GSD management models which can be used to reduce the rising costs typical of a global environment; (2) *Establishing of initial rules for resolving conflicts* [82], [89], [95], which means, determining a set of initial rules and a contract which is sufficiently flexible to avoid potential conflicts during project development; and (3) *Tools with communication features that support collaborative environments,*

since communication technologies, such as video conferencing, e-mail, and groupware tools that support virtual collaborative work, increasingly substitute the need for physical presence [107].

Lack of experience in GSD [15], [100], [101] refers to the risk involved when the client lacks the necessary knowledge about GSD project management. No specific safeguards to solve this risk have been found in explicit form in the SRL.

Lack of coordination [5], [16], [17], [44], [52], [79], [95] occurs as result of a loss of interaction and intensive communication between companies in GSD, thus causing work repetition and project delays. Several safeguards with which to manage this canonical risk in a global environment have been found in the literature: (1) *Common processes will be suitable for the GSD model* [50], [71], suggesting a common process for requirements analysis and documentation; (2) *Common tools will be established for every work group* [7], [50], [70], [82], proposing the common use of tools for online group work; (3) *Ensure there is a good team* [82], [112] with language skills that leverage communication and cooperation between all the parties involved; (4) *Increase in collaboration and coordination between companies* [5], [7], [44], [52], [95], [102], which can be achieved through frequent meetings, including formal and informal discussions; (5) *Well-defined agile leadership hierarchy* [17], [82], which, together with frequent communication, promotes better coordination; (6) *Knowledge sharing* [6], [7], [44], [71], [88], [107], [108]; this can be improved, for example, by placing team members in other groups: the sharing of culture, people and products can greatly improve the coordination of the project; and finally (7) *The moderator will chair requirements negotiation discussions to solve conflicts and to reach shared solutions* [116].

Lack of leadership [17], [79], [100]. There may be a lack of clear leadership in a globalized environment, thus causing a lack of responsibilities. Problems between analysts and project managers and even a lack of support from senior management can appear, all of which may cause delays, or even the cancelation of the project. The canonical safeguard that we have associated with this risk is *Efficient, well-defined leadership hierarchy* [17], [82]: there is a need to establish a clear leadership hierarchy [17], while the working method provides all groups with the flexibility and adaptability needed [82].

Processes that are not well-defined [15], [16], [50], [75], [82], [88], [100], [107], [122]–[124]. Attempts are frequently made to implement in GSD process models that are typical of traditional environments. The case of different groups using completely different software development process models also occurs frequently [16], [82]; this causes delays in the project, as well as increases in costs. Inconsistencies in work practice, with the offshore company employing different work practices, leads to incomplete information on the status of the project [82]. In addition, documented processes may not have been implemented in GSD [50]; poor control of changes, together with changes not properly managed in the

initial set of requirements, may therefore result in hostile situations concerning the final price or penalty clauses [100]. As a canonical safeguard for this risk, we have identified (1) *Common processes will be suitable for the GSD model* [50], [71]: all the groups involved in GSD should use common processes for analysis, documentation and change management in requirements [50]; and (2) *A common requirements notation will be chosen*. Laurent et al. [124] propose the use of the so-called *Collaborative Global Requirements Engineering Notation (CGREN)*, which allows project managers to plan, analyze, and optimize their Global RE processes, so that they can better understand their existing processes, identify weaknesses and problems, and establish essential processes and infrastructures.

Conflict between client and vendor [86], [100], [122], [125] may have several causes: for instance, there is denial of liability by a client, or the client may refuse to give in, or there is a lack of client involvement when developing requirements. Although these problems are also typical of colocalized environments, they are worse in GSD because of distance, language and use of inappropriate communication tools. Canonical safeguards to deal with this canonical risk are: (1) *New methodologies to support GSD* [105], [126], [127]. Agile methods in GSD -see e.g. [75], [92] can help to handle client-vendor relationships, but establishing initial agreements with clients and the monitoring of the development process are still two important issues [128]. The examination of the performance of agile RE in outsourced projects with distributed teams and customers remains challenging [88], and issues related to agile software development are the ones practitioners would most like to be researched [128]. Attention should also be paid to coaching globally distributed software development projects to adopt agile methods [129]; and (2) *Tools with support for iterative planning and negotiation* [78], [105], [126], [127], [130], [131] are also important in enabling those new methodologies.

Differences between the process models used [101], [122], since it is usual to find clients with no work standards or with process models that are entirely incompatible with those of the vendor company. Canonical safeguards found for this canonical risk are the following: (1) *Common processes will be suitable for the GSD model* [50], [71]; and (2) *Common tools will be established for every work group* [7], [50], [70], [82].

5) CLIENT-VENDOR DISTANCE

If client-vendor relationships during RE involve difficulties in traditional developments [132], then even more complications may appear in globally distributed environments, with new canonical risks that arise from the characteristics of GSD, such as the reduction in communication and the increase in distance discussed above. This is an important concern, as the way the client perceives company performance is even more significant than how the company perceives its performance [98]. The canonical risks identified in this concern are listed below:

Difficulty of extracting relevant information from stakeholders [5], [16], which establishes that in a global environment the elicitation of requirements that are clearly identified and shared by all the parties involved is more complex. No specific safeguards to this risk have been found in the literature.

Difficulty of identifying the key users [71], [118], concerning the difficulty involved in accessing the key or main users in GSD. Not being in the same workplace as the analysts implies that the tasks carried out by each user cannot be observed, thus causing incomplete or wrong requirements. Appropriate safeguards for this risk have not been found in the literature.

High number of users as requirements sources [50], [133]: it is common to find both a large number of users and a large number of requirements in GSD; it is even possible that thousands of relevant stakeholders in the authoring of requirements may be remotely distributed. This open, global, inclusive elicitation method would make it easier to capture a complete set of requirements, would facilitate the exploration of options in greater depth, and thereby encourage the consideration of more perspectives in the product [133]. One safeguard has been identified in the literature: *Choose an appropriate, automated prioritization requirements technique* [93], [133]: an appropriate, automated classification requirements technique should be used, such as Naïve Bayes classifier, PLSA clustering model, Analytic Hierarchy Process, Cumulative Voting, or Likert Scale Technique.

Lack of consideration of the expectations of end-users [83], [100], due to the fact that users are not in contact with developers; this can cause a lack of expectations and may even result in a situation in which the software is not accepted by the clients. We believe that the canonical safeguard *New methodologies to support GSD* [105], [126], [127], which is related to agile method adoption, can be useful here.

6) TOOLS WHICH SUPPORT THE PROCESS

Tools supporting Global RE process models must be analyzed in the quest to identify canonical risks that may occur in GSD, as well as to discover the best means of managing them. The following are the two canonical risks identified in the literature:

Lack of suitable tools for GSD [7], [50], [71], [78], [99], [134]–[138]. There is a need for tools which leverage the requirements specification process in GSD. Furthermore, the analysis and negotiation of the requirements is not easy in GSD, since the stakeholders are usually unable to access all the relevant information about the requirements, and cannot easily manage requirements traceability. Besides, a widely used way to achieve better collaboration in heterogeneous contexts is to establish an integrated, standardized technological infrastructure. Ethnographically-inspired studies, on the other hand, have challenged such a perspective, pointing out that generic technology does not fit in local contexts and that it needs to be worked-around [138]. This canonical risk could be addressed by

means of the following canonical safeguards: (1) *Tools with communication features that support collaborative environments* [77], [107], [134], [136]: the use of collaborative tools that enhance knowledge sharing and coordination of distributed teams, such as videoconferences, instant messaging tools, semantic wikis, and forums; (2) *Tools with collaborative and coordination features that support distributed environments* [72], [74], [99], [104], [112], [119], [134], [135], [137]–[146]; (3) *Tools with support for iterative planning and negotiation* [78], [105], [126], [127], [130], [131]: new development methodologies, such as agile methods, can adopt tailored tools that are well suited for negotiating requirements in global environments [126]; and finally (4) *Common tools will be established for every work group* [7], [50], [70], [82].

Tools not employed due to lack of performance and security [50]. Available tools are often not used in GSD, either because they do not meet performance needs or because they do not incorporate the appropriate security mechanisms to allow them to be used on the Internet. No safeguards with which to solve this risk have been found in explicit form in the literature.

REFERENCES

- [1] D. Damian, F. Lanubile, and H. L. Oppenheimer, "Addressing the challenges of software industry globalization: The workshop on global software development," in *Proc. 25th Int. Conf. Softw. Eng. (ICSE)*, Washington, DC, USA, 2003, pp. 793–794.
- [2] J. Binder, *Global Project Management: Communication, Collaboration and Management Across Borders*. Farnham, U.K.: Gower, 2007.
- [3] R. Sangwan, M. Bass, N. Mullick, D. J. Paulish, and J. Kazmeier, *Global Software Development Handbook*. Boca Raton, FL, USA: Taylor & Francis, 2007.
- [4] E. Carmel, *Global Software Teams: Collaborating Across Borders and Time Zones*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [5] K. Schmid, "Challenges and solutions in global requirements engineering—A literature survey," in *Proc. 6th Int. Conf. Softw. Qual. Model-Based Approaches Adv. Softw. Syst. Eng.*, Vienna, Austria, D. Winkler, S. Biffl, J. Bergsmann, Eds. Cham, Switzerland: Springer, Jan. 2014, pp. 85–99.
- [6] P. W. L. Vlaar, P. C. van Fenema, and V. Tiwari, "Cocreating understanding and value in distributed work: How members of onsite and offshore vendor teams give, make, demand, and break sense," *MIS Quart.*, vol. 32, pp. 227–255, Jun. 2008.
- [7] P. Parviainen and M. Tihinen, "Knowledge-related challenges and solutions in GSD," *Expert Syst.*, vol. 31, no. 3, pp. 253–266, 2014.
- [8] C. Ebert, *Global Software and IT: A Guide to Distributed Development, Projects, and Outsourcing*, 1st ed. Hoboken, NJ, USA: Wiley, 2011.
- [9] K. Herrmann, "Release planning in distributed projects," presented at the 1st Int. Global Requirements Eng. Workshop (GREW), Munich, Germany, 2007.
- [10] C. Ebert, C. H. Parro, R. Suttels, and H. Kolarczyk, "Improving validation activities in a global software development," presented at the 23rd Int. Conf. Softw. Eng., Toronto, ON, Canada, 2001.
- [11] M. Lormans, H. Van Dijk, A. Van Deursen, N. Eric, and A. de Zeeuw, "Managing evolving requirements in an outsourcing context: An industrial experience report," in *Proc. 7th Int. Workshop Princ. Softw. Evol.*, 2004, pp. 149–158.
- [12] H. K. Edwards and V. Sridhar, "Analysis of software requirements engineering exercises in a global virtual team setup," *J. Global Inf. Manage.*, vol. 13, no. 2, pp. 21–41, 2005.
- [13] B. H. C. Cheng and J. M. Atlee, "Research directions in requirements engineering," presented at the Future Softw. Eng., 2007.

- [14] D. Damian, J. Chisan, P. Allen, and B. Corrie, "Awareness meets requirements management: Awareness needs in global software development," presented at the Int. Workshop Global Softw. Develop., Portland, OR, USA, 2003.
- [15] R. Prikladnicki, J. L. N. Audy, D. Damian, and T. C. de Oliveira, "Distributed software development: Practices and challenges in different business strategies of offshoring and onshoring," presented at the Int. Conf. Global Softw. Eng., 2007.
- [16] D. Damian, "Stakeholders in global requirements engineering: Lessons learned from practice," *IEEE Softw.*, vol. 24, no. 2, pp. 21–27, Mar. 2007.
- [17] B. Berenbach, "Impact of organizational structure on distributed requirements engineering processes: Lessons learned," presented at the 1st Int. Workshop Global Softw. Develop. Practitioner, Shanghai, China, 2006.
- [18] J. D. Herbsleb, "Global software engineering: The future of socio-technical coordination," presented at the Future Softw. Eng., 2007.
- [19] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Dept. Comput. Sci., Univ. Durham, Durham, U.K., EBSE Tech. Rep. EBSE-2007-01, 2007.
- [20] B. A. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering—A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.
- [21] R. Matavire and I. Brown, "Profiling grounded theory approaches in information systems research," *Eur. J. Inf. Syst.*, vol. 22, no. 1, pp. 119–129, 2013.
- [22] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering*. Berlin, Germany: Springer, 2012.
- [23] J. M. Verner, O. P. Brereton, B. A. Kitchenham, M. Turner, and M. Niazi, "Risks and risk mitigation in global software development: A tertiary study," *Inf. Softw. Technol.*, vol. 56, no. 1, pp. 54–78, 2014.
- [24] J. F. Wolfswinkel, E. Furtmueller, and C. Wilderom, "Using grounded theory as a method for rigorously reviewing literature," *Eur. J. Inf. Syst.*, vol. 22, no. 1, pp. 45–55, 2013.
- [25] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quart.*, vol. 26, no. 2, pp. 14–23, 2002.
- [26] M. Wiesche, M. C. Jurisch, P. W. Yetton, and H. Krcmar, "Grounded theory methodology in information systems research," *MIS Quart.*, vol. 41, no. 3, pp. 685–701, 2017.
- [27] M. C. Lacity, S. A. Khan, and L. P. Willcocks, "A review of the IT outsourcing literature: Insights for practice," *J. Strategic Inf. Syst.*, vol. 18, no. 3, pp. 130–146, 2009.
- [28] L. Willcocks and H. Margetts, "Risk assessment and information systems," *Eur. J. Inf. Syst.*, vol. 3, no. 2, pp. 127–138, 1994.
- [29] L. P. Willcocks and M. C. Lacity, "IT outsourcing in insurance services: Risk, creative contracting and business advantage," *Inf. Syst. J.*, vol. 9, no. 3, pp. 163–180, 1999.
- [30] W. Caelli and D. Longley, *Information Security for Managers*, 1st ed. Basingstoke, U.K.: Palgrave Macmillan, 1989.
- [31] T. Finne, "Information systems risk management: Key concepts and business processes," *Comput. Secur.*, vol. 19, no. 3, pp. 234–242, 2000.
- [32] B. W. Boehm, "Software risk management: Principles and practices," *IEEE Softw.*, vol. 8, no. 1, pp. 32–41, Jan. 1991.
- [33] S. R. Nidumolu, "Standardization, requirements uncertainty and software project performance," *Inf. Manage.*, vol. 31, pp. 135–150, Dec. 1996.
- [34] J. Biolchini, P. G. Mian, A. C. C. Natali, and G. H. Travassos, "Systematic review in software engineering," Dept. Syst. Eng. Comput. Sci., COPPE/UF RJ, Rio de Janeiro, Brazil, Tech. Rep. TR-ES 679/05, May 2005.
- [35] F. Calefato, D. Gendarmi, and F. Lanubile, "Investigating the use of tags in collaborative development environments: A replicated study," presented at the ACM-IEEE Int. Symp. Empirical Softw. Eng. Meas., Bolzano-Bozen, Italy, 2010.
- [36] H. Ossher et al., "Using tagging to identify and organize concerns during pre-requirements analysis," presented at the ICSE Workshop Aspect-Oriented Requirements Eng. Archit. Design, 2009.
- [37] P. Mika, *Social Networks and the Semantic Web*, 1st ed. New York, NY, USA: Springer, 2007.
- [38] B. S. Everitt, S. Landau, and M. Leese, *Cluster Analysis*. Hoboken, NJ, USA: Wiley, 2009.
- [39] E. Mooi and M. Sarstedt, "Cluster analysis," in *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*, 1st ed. Berlin, Germany: Springer-Verlag, 2011.
- [40] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," *J. Syst. Softw.*, vol. 80, no. 4, pp. 571–583, 2007.
- [41] M. C. Lacity, S. Khan, A. Yan, and L. P. Willcocks, "A review of the IT outsourcing empirical literature and future research directions," *J. Inf. Technol.*, vol. 25, no. 4, pp. 395–433, 2010.
- [42] F. Elberzhager, J. Münch, and V. T. N. Nha, "A systematic mapping study on the combination of static and dynamic quality assurance techniques," *Inf. Softw. Technol.*, vol. 54, no. 1, pp. 1–15, 2012.
- [43] A. Ampatzoglou, S. Charalampidou, and I. Stamelos, "Research state of the art on GoF design patterns: A mapping study," *J. Syst. Softw.*, vol. 86, pp. 1945–1964, Jul. 2013.
- [44] A. Ahmad and H. Khan, "The importance of knowledge management practices in overcoming the global software engineering challenges in requirements understanding," M.S. thesis, School Eng., Blekinge Inst. Technol., Karlskrona, Sweden, 2008.
- [45] N. B. Ali and M. Usman, "Reliability of search in systematic reviews: Towards a quality assessment framework for the automated-search strategy," *Inf. Softw. Technol.*, vol. 99, pp. 133–147, Jul. 2018.
- [46] B. Ulziit, Z. A. Warraich, C. Gencel, and K. Petersen, "A conceptual framework of challenges and solutions for managing global software maintenance," *J. Softw. Evol. Process*, vol. 27, no. 10, pp. 763–792, 2015.
- [47] B. A. Aubert, H. Barki, M. Patry, and V. Roy, "A multi-level, multi-theory perspective of information technology implementation," *Inf. Syst. J.*, vol. 18, no. 1, pp. 45–72, 2008.
- [48] C. E. H. Chua, W.-K. Lim, C. Soh, and S. K. Sia, "Client strategies in vendor transition: A threat balancing perspective," *J. Strategic Inf. Syst.*, vol. 21, no. 1, pp. 72–83, 2012.
- [49] D. Zowghi and D. E. Damian, "An insight into the interplay between culture, conflict and distance in globally distributed requirements negotiations," presented at the 36th Hawaii Int. Conf. Syst. Sci., Island, HI, USA, 2003.
- [50] T. Illes-Seifert, A. Herrmann, M. Geisser, and T. Hildenbrand, "The challenges of distributed software engineering and requirements engineering: Results of an online survey," presented at the 1st Int. Global Requirements Eng. Workshop, Munich, Germany, 2007.
- [51] M. Helén, "Challenges in multi-site and multi-cultural globally distributed software development," M.S. thesis, Dept. Inf. Syst. Sci., Univ. Jyväskylä, Jyväskylä, Finland, 2004.
- [52] I. Kwan, D. Damian, and S. Marczak, "The effects of distance, experience, and communication structure on requirements awareness in two distributed industrial software projects," presented at the 1st Int. Global Requirements Eng. Workshop, Munich, Germany, 2007.
- [53] T. Ebling, J. L. N. Audy, and R. Prikladnicki, "A systematic literature review of requirements engineering in distributed software development environments," in *Proc. 11th Int. Conf. Enterprise Inf. Syst. (ICEIS)*, Milan, Italy, 2009, pp. 363–366.
- [54] A. A. Khan, S. Basri, F. E. Amin, U. Teknologi, T. Perak, and I. Studies, "Communication risks and best practices in global software development during requirements change management: A systematic literature review protocol," *Res. J. Appl. Sci., Eng. Technol.*, vol. 6, no. 19, pp. 3514–3519, 2013.
- [55] J. Noll, S. Beecham, and I. Richardson, "Global software development and collaboration: Barriers and solutions," *ACM Inroads*, vol. 1, no. 3, pp. 66–78, 2011.
- [56] R. Jabangwe and I. Nurdiani, "Global software development challenges and mitigation strategies: A systematic review and survey results," M.S. thesis, Blekinge Inst. Technol., Karlskrona, Sweden, 2010.
- [57] F. Q. B. da Silva, R. Prikladnicki, A. C. C. França, C. V. F. Monteiro, C. Costa, and R. Rocha, "An evidence-based model of distributed software development project management: Results from a systematic mapping study," *J. Softw., Evol. Process*, vol. 24, no. 6, pp. 625–642, 2012.
- [58] F. Q. B. da Silva, C. Costa, A. C. C. França, and R. Prikladnicki, "Challenges and solutions in distributed software development project management: A systematic literature review," presented at the 5th IEEE Int. Conf. Global Softw. Eng., Princeton, NJ, USA, Aug. 2010.
- [59] M. Jiménez, M. Piatini, and A. Vizcaíno, "Challenges and improvements in distributed software development: A systematic review," *Adv. Softw. Eng.*, vol. 2009, Jan. 2009, Art. no. 3.

- [60] S. U. Khan, M. Niazi, and R. Ahmad, "Barriers in the selection of off-shore software development outsourcing vendors: An exploratory study using a systematic literature review," *Inf. Softw. Technol.*, vol. 53, no. 7, pp. 693–706, 2010.
- [61] D. Šmite, C. Wohlin, T. Gorschek, and R. Feldt, "Empirical evidence in global software engineering: A systematic review," *Empirical Softw. Eng.*, vol. 15, no. 1, pp. 91–118, 2010.
- [62] C. Costa, C. Cunha, R. Rocha, A. C. C. França, F. Q. B. da Silva, and R. Prikladnicki, "Models and tools for managing distributed software development: A systematic literature," presented at the 14th Int. Conf. Eval. Assessment Softw. Eng., Keele, U.K.: Keele Univ., 2010.
- [63] R. Prikladnicki and J. L. N. Audy, "Process models in the practice of distributed software development: A systematic review of the literature," *Inf. Softw. Technol.*, vol. 52, no. 8, pp. 779–791, 2010.
- [64] S. Nidhra, M. Yanamadala, W. Afzal, and R. Torkar, "Knowledge transfer challenges and mitigation strategies in global software development—A systematic literature review and industrial validation," *Int. J. Inf. Manage.*, vol. 33, no. 2, pp. 333–355, 2013.
- [65] A. Lopez, J. Nicolas, and A. Toval, "Risks and safeguards for the requirements engineering process in global software development," in *4th IEEE Int. Conf. Global Softw. Eng. (ICGSE)*, Limerick, Ireland, 2009, pp. 394–399.
- [66] P. Louridas, "Using wikis in software development," *IEEE Softw.*, vol. 23, no. 2, pp. 88–91, Mar. 2006.
- [67] CMMI Product Team. (2007). *CMMI for Acquisition, Version 1.2 (CMU/SEI-2007-TR-017)*. [Online]. Available: <http://www.sei.cmu.edu/library/abstracts/reports/07tr017.cfm>
- [68] D. E. Damian, "The study of requirements engineering in global software development: As challenging as important," presented at the 22rd Int. Conf. Softw. Eng., Sydney, FL, USA, 2002.
- [69] R. Prikladnicki, R. Evaristo, K. Gallagher, L. T. Lopes, and J. L. N. Audy, "The role of culture in interpreting qualitative data: Methodological issues in an exploratory study of cross-cultural distributed software development," presented at the 13th Annu. Cross-Cultural Meeting Inf. Syst. (ICIS), Las Vegas, NV, USA, 2005.
- [70] V. Clerc, "Towards architectural knowledge management practices for global software development," presented at the 3rd Int. Workshop Sharing Reusing Archit. Knowl., Leipzig, Germany, 2008.
- [71] S. Betz, A. Oberweis, and R. Stephan, "Knowledge transfer in offshore outsourcing software development projects: An analysis of the challenges and solutions from German clients," *Expert Syst.*, vol. 31, no. 3, pp. 282–297, 2012.
- [72] C. Sillaber and R. Breu, "The impact of knowledge sharing platforms in distributed requirements engineering scenarios: A systematic review," in *Proc. 8th Int. Conf. Knowl. Manage. Organizations, Social Big Data Comput. Knowl. Manage.*, L. Uden, L. Wang, J. C. Rodríguez, H.-C. Yang, and I. H. Ting, Eds. Dordrecht, The Netherlands: Springer, 2014, pp. 579–591.
- [73] D. Damian, "The use of a multimedia group support system for distributed software requirements meetings," in *Proc. HICCS*, 2002, pp. 1–9.
- [74] B. Decker, E. Ras, J. Rech, P. Jaubert, and M. Rieth, "Wiki-based stakeholder participation in requirements engineering," *IEEE Softw.*, vol. 24, no. 2, pp. 28–35, Mar. 2007.
- [75] T. O. A. Lehtinen, R. Virtanen, V. T. Heikkilä, and J. Itkonen, "Why the development outcome does not meet the product owners' expectations?" in *Proc. 16th Int. Conf. Agile Processes Softw. Eng. Extreme Program. (XP)*, Helsinki, Finland, C. Lassenius, T. Dingsøyr, and M. Paasivaara, Eds. Cham, Switzerland: Springer, May 2015, pp. 93–104.
- [76] F. Calefato, D. Damian, and F. Lanubile, "Computer-mediated communication to support distributed requirements elicitation and negotiations tasks," *Empirical Softw. Eng.*, vol. 17, no. 6, pp. 640–674, 2012.
- [77] D. Damian, F. Lanubile, and T. Mallardo, "On the need for mixed media in distributed requirements negotiations," *IEEE Trans. Softw. Eng.*, vol. 34, no. 1, pp. 116–132, Jan. 2008.
- [78] N. Dhruv, S. Varadarajan, A. Monica, and M. Amit, "Project quality of off-shore virtual teams engaged in software requirements analysis: An exploratory comparative study," *J. Global Inf. Manage.*, vol. 16, no. 4, pp. 24–45, 2008.
- [79] D. C. Gumm, "A model of requirements engineering at organizational interfaces: An empirical study on distributed requirements engineering," presented at the 1st Int. Global Requirements Eng. Workshop, Munich, Germany, 2007.
- [80] S. I. Hashmi, F. Ishikawa, and I. Richardson, "A communication process for global requirements engineering," presented at the Int. Conf. Softw. Syst. Process, San Francisco, CA, USA, 2013.
- [81] K. Stapel and K. Schneider, "Managing knowledge on communication and information flow in global software projects," *Expert Syst.*, vol. 31, no. 3, pp. 234–252, 2012.
- [82] N. B. Moe and D. Šmite, "Understanding a lack of trust in global software teams: A multiple-case study," *Softw. Process, Improvement Pract.*, vol. 13, no. 3, pp. 217–231, 2008.
- [83] D. Šmite, "Requirements management in distributed projects," *J. Universal Knowl. Manage.*, vol. 1, no. 2, pp. 69–76, 2006.
- [84] D. Damian, "An empirical study of requirements engineering in distributed software projects: Is distance negotiation more effective?" presented at the 8th Asia-Pacific Softw. Eng. Conf., 2001.
- [85] C. Rosenk, H. Vranešić, and R. Holten, "Boundary interactions and motors of change in requirements elicitation: A dynamic perspective on knowledge sharing," *J. Assoc. Inf. Syst.*, vol. 15, no. 6, pp. 306–345, 2014.
- [86] D. E. Damian and D. Zowghi, "The impact of stakeholders' geographical distribution on managing requirements in a multi-site organization," presented at the 10th Anniversary IEEE Joint Int. Conf. Requirements Eng., Essen, Germany, Sep. 2002.
- [87] M. Heindl and S. Biffl, "Risk management with enhanced tracing of requirements rationale in highly distributed projects," in *Proc. 1st Int. Workshop Global Softw. Develop. Practitioner (GSD)*, Shanghai, China, 2006, pp. 20–26.
- [88] I. Inayat, S. S. Salim, S. Marczak, M. Daneva, and S. Shamshirband, "A systematic literature review on agile requirements engineering practices and challenges," *Comput. Hum. Behav.*, vol. 51, pp. 915–929, Oct. 2015.
- [89] M. Nordio, R. Mitin, B. Meyer, C. Ghezzi, E. Di Nitto, and G. Tamburrelli, "The role of contracts in distributed development," in *Proc. 3rd Int. Conf. Softw. Eng. Approaches Offshore Outsourced Develop. (SEAFOOD)*, Zürich, Switzerland, O. Gotel, M. Joseph and B. Meyer, Eds. Berlin, Germany: Springer, Jul. 2009, pp. 117–129.
- [90] J. L. F. Alemán, J. M. Carrillo-de-Gea, J. V. Meca, J. N. Ros, A. Toval, and A. Idri, "Effects of using requirements catalogs on effectiveness and productivity of requirements specification in a software project management course," *IEEE Trans. Educ.*, vol. 59, no. 2, pp. 105–118, May 2016.
- [91] J. M. C. de Gea, J. Nicolás, J. L. F. Alemán, A. Toval, S. Ouhbi, and A. Idri, "Co-located and distributed natural-language requirements specification: Traditional versus reuse-based techniques," *J. Software, Evol. Process*, vol. 28, no. 3, pp. 205–227, 2016.
- [92] S. Ghobadi and L. Mathiassen, "Perceived barriers to effective knowledge sharing in agile software teams," *Inf. Syst. J.*, vol. 26, no. 2, pp. 95–125, 2016.
- [93] S. Scheibmayr, T. Hildenbrand, and M. Geisser, "An experimental, tool-based evaluation of requirements prioritization techniques in distributed settings," in *Proc. AMCIS*, 2009, p. 735.
- [94] M. Daneva et al., "Agile requirements prioritization in large-scale outsourced system projects: An empirical study," *J. Syst. Softw.*, vol. 86, pp. 1333–1353, May 2013.
- [95] S. Overhage, O. Skroch, and K. Turowski, "A method to evaluate the suitability of requirements specifications for offshore projects," *Bus. Inf. Syst. Eng.*, vol. 2, no. 3, pp. 155–164, 2010.
- [96] M. W. Bhatti and A. Ahsan, "Global software development: An exploratory study of challenges of globalization, HRM practices and process improvement," *Rev. Managerial Sci.*, vol. 10, no. 4, pp. 649–682, 2015.
- [97] I. van de Weerd, S. Brinkkemper, and J. Versendaal, "Incremental method evolution in global software product management: A retrospective case study," *Inf. Softw. Technol.*, vol. 52, no. 7, pp. 720–732, 2010.
- [98] T. Tunkelo, A.-P. Hameri, and Y. Pigneur, "Improving globally distributed software development and support processes—A workflow view," *J. Softw., Evol. Process*, vol. 25, no. 12, pp. 1305–1324, 2013.
- [99] M. Heindl, F. Reinisch, and S. Biffl, "Requirements management infrastructures in global software development," presented at the Int. Workshop Tool Support Requirements Manage. Distrib. Projects, Munich, Germany, 2007.
- [100] C. L. Iacovou and R. Nakatsu, "A risk profile of offshore-outsourced development projects," *Commun. ACM*, vol. 51, no. 6, pp. 89–94, 2008.
- [101] R. Prikladnicki, J. Audy, and R. Evaristo, "Requirements management in global software development: Preliminary findings from a case study in a SW-CMM context," presented at the Int. Workshop Global Softw. Develop., Portland, OR, USA, 2003.

- [102] I. Inayat and S. S. Salim, "A framework to study requirements-driven collaboration among agile teams: Findings from two case studies," *Comput. Hum. Behav.*, vol. 51, pp. 1367–1379, Oct. 2015.
- [103] S. Islam, S. H. Houb, D. Mendez-Fernandez, and M. M. A. Joarder, "Offshore-outsourced software development risk management model," presented at the 12th Int. Conf. Comput. Inf. Technol., Dhaka, Bangladesh, 2009.
- [104] O. Uenal, N. Riegel, S. Weber, and J. Doerr, "Using enhanced wiki-based solutions for managing requirements," presented at the 1st Int. Workshop Manag. Requirements Knowl., Barcelona, Spain, 2008.
- [105] A. Morali and R. Wieringa, "Risk-based confidentiality requirements specification for outsourced IT systems," presented at the 18th IEEE Int. Requirements Eng. Conf., Sydney, NSW, Australia, Sep./Oct. 2010.
- [106] H. Bruun and S. Sierla, "Distributed problem solving in software development: The case of an automation project," *Social Stud. Sci.*, vol. 38, pp. 133–158, Feb. 2008.
- [107] J. Dibbern, J. Winkler, and A. Heinzl, "Explaining variations in client extra costs between software projects offshored to India," *MIS Quart.*, vol. 32, no. 2, pp. 333–366, 2008.
- [108] V. Vijayakumar, R. Gey, and E. Wende, "Storytelling—A method to start knowledge transfer in offshore software development teams," presented at the 9th Eur. Conf. Knowl. Manage. (ECKM), Southampton, U.K., 2008.
- [109] J. K. Winkler, J. Dibbern, and A. Heinzl, "The impact of cultural differences in offshore outsourcing—Case study results from German–Indian application development projects," *Inf. Syst. Frontiers*, vol. 10, no. 2, pp. 243–258, 2008.
- [110] J. Dibbe, W. W. Chin, and A. Heinzl, "Systemic determinants of the information systems outsourcing decision: A comparative study of German and United States firms," *J. Assoc. Inf. Syst.*, vol. 13, no. 6, pp. 466–497, 2012.
- [111] M. Ruano-Mayoral, C. Casado-Lumbreras, H. Garbarino-Alberti, and S. Misra, "Methodological framework for the allocation of work packages in global software development," *J. Softw., Evol. Process*, vol. 26, no. 6, pp. 476–487, 2014.
- [112] P. Holtkamp, J. P. P. Jokinen, and J. M. Pawlowski, "Soft competency requirements in requirements engineering, software design, implementation, and testing," *J. Syst. Softw.*, vol. 101, pp. 136–146, Mar. 2015.
- [113] M. Crofts, R. Smith, and B. Fraunholz, "Global software development: The next RE frontier?" presented at the Austral. Workshop Requirements Eng., Adelaide, SA, USA, 2004, pp. 1–12.
- [114] L. Lopes, R. Prikladnicki, J. L. N. Audy, and A. Majdenbaum, "Distributed requirements specification: Minimizing the effect of geographic dispersion," presented at the 22nd Int. Conf. Softw. Eng., Tallahassee, FL, USA, 2002.
- [115] G. N. Aranda, A. Vizcaíno, and M. Piattini, "A framework to improve communication during the requirements elicitation process in GSD projects," *Requirements Eng.*, vol. 15, no. 4, pp. 397–417, 2010.
- [116] S. Deshpande, I. Richardson, V. Casey, and S. Beecham, "Culture in global software development—A weakness or strength?" presented at the 5th IEEE Int. Conf. Global Softw. Eng., Aug. 2010.
- [117] E. MacGregor, Y. Hsieh, and P. Kruchten, "Cultural patterns in software process mishaps: Incidents in global projects," presented at the Workshop Hum. Social Factors Softw. Eng., St. Louis, Missouri, 2005.
- [118] J. Hanisch and B. Corbitt, "Requirements engineering during global software development: Some impediments to the requirements engineering process—A case study," presented at the Eur. Conf. Inf. Syst., Turku, Finland, 2004.
- [119] M. Spichkova and H. Schmidt, "Requirements engineering aspects of a geographically distributed architecture," in *Proc. 10th Int. Conf. Eval. Novel Approaches Softw. Eng. (ENASE)*, 2015, pp. 276–281.
- [120] M. Spichkova, H. W. Schmidt, M. R. I. Nekvi, and N. H. Madhavji, "Structuring diverse regulatory requirements for global product development," in *Proc. IEEE 8th Int. Workshop Requirements Eng. Law (RELAW)*, Ottawa, ON, Canada, Aug. 2015, pp. 57–60.
- [121] H. Hiisilä, M. Kauppinen, and S. Kujala, "Challenges of the customer organization's requirements engineering process in the outsourced environment—A case study," in *Proc. 21th Int. Work. Conf. Requirements Eng., Found. Softw. Qual., REFSQ*, Essen, Germany, A. S. Fricker K. Schneider, Eds. Cham, Switzerland: Springer, Mar. 2015, pp. 214–229.
- [122] J. M. Bhat, M. Gupta, and S. N. Murthy, "Overcoming requirements engineering challenges: Lessons from offshore outsourcing," *IEEE Softw.*, vol. 23, no. 5, pp. 38–44, Sep./Oct. 2006.
- [123] V. Mikulovic and M. Heiss, "'How do I know what I have to do?': The role of the inquiry culture in requirements communication for distributed software development projects," presented at the 28th Int. Conf. Softw. Eng., Shanghai, China, 2006.
- [124] P. Laurent, A. Steele, J. Cleland-Huang, and P. Mäeder, "Evaluating the effectiveness of a collaborative requirements engineering modeling notation for planning globally distributed projects," in *Proc. Int. Conf. Softw. Eng. Res. Pract. (SERP)*, Athens, Greece, 2013, pp. 1–7.
- [125] V. N. Vithana, "Scrum requirements engineering practices and challenges in offshore software development," *Int. J. Comput. Appl.*, vol. 116, no. 22, pp. 43–49, 2015.
- [126] S. Jalali and C. Wohlin, "Global software engineering and agile practices: A systematic review," *J. Softw. Maintenance Evol., Res. Pract.*, 2011.
- [127] P. Laurent, P. Mader, J. Cleland-Huang, and A. Steele, "A taxonomy and visual notation for modeling globally distributed requirements engineering projects," in *Proc. 5th IEEE Int. Conf. Global Softw. Eng.*, vol. 10, Aug. 2010, pp. 35–44.
- [128] E. Papatheocharous and A. S. Andreou, "Empirical evidence and state of practice of software agile teams," *J. Softw., Evol. Process*, vol. 26, no. 9, pp. 855–866, 2014.
- [129] M. Paasivaara and C. Lassenius, "Agile coaching for global software development," *J. Softw., Evol. Process*, vol. 26, no. 4, pp. 404–418, 2014.
- [130] F. Salger, S. Sauer, G. Engels, and A. Baumann, "Knowledge transfer in global software development—Leveraging ontologies, tools and assessments," presented at the 5th IEEE Int. Conf. Global Softw. Eng., Aug. 2010.
- [131] C. Scharff, "An evolving collaborative model of working in students' global software development projects," presented at the Community Building Workshop Collaborative Teaching Globally Distrib. Softw. Develop., Honolulu, HI, USA, 2011.
- [132] S. Shuraida and H. Barki, "The influence of analyst communication in IS projects," *J. Assoc. for Inf. Syst.*, vol. 14, no. 9, pp. 482–520, 2013.
- [133] M. A. Adegoke, J. O. A. Ayeni, and F. O. Sogunle, "An empirical study of automated classification tools for informal requirements in large scale systems," *Int. J. Bus. Inf. Syst.*, vol. 19, no. 3, pp. 324–341, 2015.
- [134] F. Calefato and F. Lanubile, "Using the econference tool for synchronous distributed requirements workshops," presented at the 1st Int. Workshop Distrib. Softw. Develop., Paris, France, 2005.
- [135] F. Lanubile, "A P2P toolset for distributed requirements elicitation," presented at the 1st Int. Workshop Distrib. Softw. Develop., Paris, France, 2003.
- [136] V. Sinha, B. Sengupta, and S. Chandra, "Enabling collaboration in distributed requirements management," *IEEE Softw.*, vol. 23, no. 5, pp. 52–61, Sep. 2006.
- [137] K. Dullemond, B. van Gameren, and R. van Solingen, "Collaboration spaces for virtual software teams," *IEEE Softw.*, vol. 31, no. 6, pp. 47–53, Nov./Dec. 2014.
- [138] J. Gasparas and E. Monteiro, "Counteracting forces in implementation of IS-enabled global business processes," presented at the 17th Eur. Conf. Inf. Syst. (ECIS), Verona, Italy, 2009.
- [139] T. Fancott, P. Kamthan, and N. Shahmir, "Towards next generation requirements engineering," in *Proc. Int. Conf. Social Inform. (SocialInformatics)*, 2012, pp. 328–331.
- [140] H. Lai, R. Peng, D. Sun, F. Shao, and Y. Liu, "A survey of RE-specific wikis for distributed requirements engineering," in *Proc. 8th Int. Conf. Semantics, Knowl. Grids (SKG)*, 2012, pp. 47–55.
- [141] C. Magnusson and S.-C. Chou, "Risk and compliance management framework for outsourced global software development," presented at the 5th IEEE Int. Conf. Global Softw. Eng., Aug. 2010.
- [142] P. Liang, P. Avgeriou, and V. Clerc, "Requirements reasoning for distributed requirements analysis using semantic wiki," in *Proc. 4th IEEE Int. Conf. Global Softw. Eng.*, 2009, pp. 388–393.
- [143] C. Solis and N. Ali, "Distributed requirements elicitation using a spatial hypertext wiki," presented at the 5th IEEE Int. Conf. Global Softw. Eng., Aug. 2010.
- [144] S. Lohmann, P. Heim, S. Auer, S. Dietzold, and T. Riechert, "Semantifying requirements engineering—The softwiki approach," presented at the I-SEMANTICS, Graz, Austria, 2008.
- [145] S. Lohmann, P. Heim, and K. Lauenroth, "Web-based stakeholder participation in distributed requirements elicitation," in *Proc. 16th IEEE Int. Requirements Eng. Conf.*, Sep. 2008, pp. 323–324.
- [146] T. D. Breaux, A. I. Antón, and E. H. Spafford, "A distributed requirements management framework for legal compliance and accountability," *Comput. Secur.*, vol. 28, nos. 1–2, pp. 8–17, 2009.



JOAQUÍN NICOLÁS received the B.Sc. and Ph.D. degrees in computer science from the University of Murcia, Murcia, Spain, in 1994 and 2009, respectively.

He is currently an Associate Professor with the Department of Informatics and Systems, University of Murcia. He has authored over 30 articles published in JCR journals and international conferences (e.g., *IEEE Software*, *Information and Software Technology*, *Computer Science and Information Systems*, and the *Journal of Software: Evolution and Process*). His current research interests are in agile, continuous requirements engineering, global software development, usability, and sustainable processes.



BERNABÉ NICOLÁS received the B.Sc. degree in computer science and engineering and the M.Sc. degree in computer science and mathematics applied to science and engineering from the University of Murcia, Murcia, Spain, in 2003 and 2009, respectively.

From 2010 to 2011, he was a Research Assistant with the Department of Informatics and Systems, University of Murcia. Since 2011, he has been a Software Developer with the Development Area, TICARUM SLU, Murcia. His research interests include project management, requirements engineering, and risk management.

Mr. Nicolás is the Secretary of the Professional Association of Computer Scientists and Engineers, Murcia.



JOSÉ L. FERNÁNDEZ-ALEMÁN received the B.Sc. degree (Hons.) in computer science and the Ph.D. degree in computer science from the University of Murcia, Murcia, Spain, in 1994 and 2002, respectively.

He is currently an Associate Professor with the University of Murcia, where he is also a member of the Software Engineering Research Group. He has authored over 50 JCR papers in the areas of software engineering and requirements engineering and their application to the fields of e-health and e-learning. His current research interests are m-health and m-learning and their application to computer science, medicine, and nursing.



JUAN M. CARRILLO DE GEA received the B.Sc. degree in nursing, the five-year B.Sc. degree in computer science and engineering, the M.Sc. degree in industrial informatics, and the Ph.D. degree in software engineering from the University of Murcia, Murcia, Spain, in 2000, 2009, 2009, and 2016, respectively.

He is currently a Research Associate and an Adjunct Professor with the University of Murcia. He has authored over 20 articles in software engineering, requirements engineering, and applications in the e-health and e-learning domains in relevant journals and conferences. His current research interests include requirements engineering, software sustainability, empirical software engineering, e-health, and e-learning.



AMBROSIO TOVAL received the B.Sc. degree in mathematics from the University Complutense of Madrid, Madrid, Spain, in 1983, and the Ph.D. degree (*cum laude*) in computer science from the Technical University of Valencia, Valencia, Spain, in 1994.

He is currently a Full Professor with the University of Murcia, Spain, where he is also the Head of the Software Engineering Research Group. He has conducted a variety of research and technology transfer projects in the areas of requirements engineering processes and tools, privacy and security requirements, sustainable requirements, and applications in the e-health, e-learning, and mobile development domains. He has authored in these topics in international journals, such as *IEEE Software*, *Information and Software Technology*, *Requirements Engineering*, *Computer Standards & Interfaces*, *IET Software*, and *International Journal of Information Security*.

...