

MAPPING THE EVOLUTION OF RESEARCH ON GLOBAL SOFTWARE ENGINEERING

A Systematic Literature Review

Josiane Kroll, Jorge Luis Nicolas Audy and Rafael Prikladnicki
Pontifical Catholic University of Rio Grande do Sul (PUCRS), Porto Alegre, Brazil

Keywords: Distributed Software Development, Software Engineering, Global Software Engineering, Systematic Literature Review.

Abstract: Studies on Global Software Engineering (GSE) have increase in recent years. However, it is not clear how research on this area has evolved in terms of topics being investigated. For this reason, the purpose of this paper is to identify areas within Software Engineering (SE) about which studies related to GSE have been published and to indicate possible gaps in this recent research area. We report from a Systematic Literature Review (SLR) conducted from September to November of 2010. The main result is mapping the evolution of this research field including what has been investigated.

1 INTRODUCTION

Global Software Engineering (GSE) is a recent research field. The concept of GSE started in the 90s in consequence of the PC revolution. With the increase of the studies in GSE over the years, is it not clear how the research in this area has evolved in terms of topics being investigated.

Some studies using the Systematic Literature Review (SLR) method were conducted to contextualize the GSE research field (Smite et al. 2009; Jiménez, Piattini, and Vizcaino 2009; Ebling, Audy, and Prickladnicki, 2009). Other studies using SLR method explored aspects of the GSE, but did not discuss the evolution of research in GSE.

For this reason, this paper reports from a SLR in order to identify software engineering (SE) areas that have studies in GSE and the most discussed topics in each area. Consequently, it will allow the analysis of the evolution of the GSE research field, SE areas that have been investigated, and the topics that have to be dealt with.

The paper is organized as follows. Section 2, describes the research method that was used in this study. Section 3, presents results obtained from the SLR. Section 4, presents a critical analysis of the obtained data, highlighting issues and the evolution of GSE research field. Finally, Section 5 provides some concluding remarks.

2 SYSTEMATIC LITERATURE REVIEW PROCESS

We conducted this SLR from September to November 2010 following the procedures recommended by Kitchenham (2004). All identified studies were mapped into Knowledge Areas (KAs) from SWEBOK (Software Engineering Body of Knowledge) (Abran et al., 2004), PMBOK (Project Management Body of Knowledge) (PMBOK, 2004) or into three new KAs defined in this study. The research questions that guided this SLR were:

RQ.1: *Which KAs of SE have studies in GSE?*

RQ.2: *What are the discussed topics in each identified KA?*

The keywords that guided the search to answers to the research questions were: *global software development, global software engineering, distributed software development, and distributed software engineering.*

We searched digital libraries such as IEEE Xplore, ACM Digital Library, Wiley Inter Science, and Elsevier Science Direct. We choose the period from 1990 to 2010, because studies in GSE began to be published in the early 1990's (Prikladnick, Audy, and Shull, 2010; Smite et al., 2010). The duplicated studies were excluded from the research. In total we found 981 studies, out of which 602 were selected for analysis.

3 FINDINGS

3.1 Software Engineering Areas with Studies in GSE (RQ.1)

To answer RQ.1, the KAs defined on SWEBOK and PMBOK were used to classify the studies. However, during classification process, we observed that both SWEBOK and PMBOK did not have KAs that could be used to classify some of the studies. This way, we defined three additional KAs:

- General Area: holds studies that report results from research in GSE that are not in KAs from the SWEBOK nor the PMBOK.
- Position Paper: studies that report an author’s point of view on a given issue in GSE.
- Systematic Review: studies of the literature using the systematic review method. Studies in this area can also be classified in KAs from the SWEBOK or PMBOK.

In Table 1 we present the number of studies found for each KAs of the SWEBOK.

Table 1: Number of Studies in the SWEBOK KAs.

SWEBOK KAs	Number of Studies
Software Engineering Processes	87
Methods and Tools of SE	69
Software Project	51
Software Engineering Management	46
Software Requirements	33
Software Construction	32
Software Quality	15
Software Maintenance	13
Software Testing	11
Software Configuration Management	4

As for the PMBOK, the studies are distributed in seven KAs, as presented in Table 2.

Table 2: Number of Studies in the PMBOK KAs.

PMBOK KAs	Number of Studies
Project Communications Management	51
Project Human Resource Management	40
Project Quality Management	11
Project Risk Management	11
Project Integration Management	9
Project Scope Management	4
Project Cost Management	3
Project Time Management	0
Project Procurement Management	0

Some studies could not be classified in any SWEBOK or PMBOK KAs. However, these studies

were mapped into the new KAs defined in this research (Table 3).

Table 3: Number of Studies in the New KAs.

New KAs	Total
General Area	99
Systematic Review	8
Position Paper	5

3.2 Research Topics Identified (RQ.2)

It is important to observe that each study could address more than on topic. Details of the identified topics are described in following subsections.

3.2.1 Research Topics in SWEBOK KAs

In Table 4 we present the main topics found for the Software Requirements KA.

Table 4: Topics of the Software Requirements KA.

Topics	Area%	Total%
Approaches to elicitation and trading requirements; Use of social networks, semantics data and wiki;	18.2	1.0
Tools for systems requirements management;	12.1	0.7
Simulators for training and learning requirements elicitation;	9.1	0.5
Process support system; Aspects of culture and distance; Challenges in the negotiation of requirements; Requirements extraction techniques;	6.1	0.3
Notations for requirements modeling; Validation of requirements; Aspects of communication in process of requirements extracting; Text-based communication for elicitation of requirements; UML models of requirements; Frameworks; Models; Volatile Requirements.	3.0	0.2

The Software Project KA had 51 studies and is the third KA with the largest percentage of studies found (8.5%). The most discussed topics in this KA are presented in Table 5. The Software Construction KA represents 5.3% of the studies found (Table 6). The Software Testing KA is the second least explored area by studies in GSE. It concentrates 1.8% of the total of studies (Table 7). The Software Maintenance KA is also little explored in GSE research, with only 1.7% of the studies (Table 8).

Table 5: Topics of the Software Project KA.

Topics	Are%	Total%
Software architecture and structure	33.3	2.8
Development tools	19.6	1.7
Development methodologies	9.8	0.8
Construction of models, UML models and prototypes; Environment management processes; Design patterns; Characteristics of the structure of GSE; Frameworks;	5.9	0.5
Inner source software development; Taxonomies for identification of the GSE dimensions; Wiki; Development via mobile.	2.0	0.2

Table 6: Topics of the Software Construction KA.

Topics	Areas%	Total%
Project theories and models;	28.1	1.5
Open source development; Implementation tools for virtual environments; Development Approaches;	12.5	0.7
Languages;	9.4	0.5
Methodologies; Managing artifacts; Tools; Using components; Patterns; Business models;	6.3	0.3
Tasks allocation; Agile methodologies; Platforms for object oriented languages; CORBA based distributed components; Development testing; Social networks.	3.1	0.2

Table 7: Topics of the Software Testing KA.

Topics	Area%	Total%
Processes for defect detection and simulation;	27.3	0.5
Experiences across teams; Effectiveness test models;	18.2	0.3
Performance assessment methodologies of applications; Allocation of developers and testers to perform tests; Training tutorials; Internet-based approaches to testing; Verification and validation; Tools.	9.1	0.2

Table 8: Topics of the Software Maintenance KA.

Topics	Area%	Total%
Sources of error in products and processes; Management mechanisms;	30.8	0.7
Cost estimation;	23.1	0.5
Collaborative work; Models.	7.7	0.2

In the Software Configuration Management KA, 4 studies and 3 topics were found. This area corresponds to 0.6% of the studies found. This KA has the least number of studies in GSE. Table 9 presents this data.

Table 9: Topics of the Software Configuration Management KA.

Topics	Area%	Total%
Managing documentation versions; Tools coordination;	50.0	0.3
Security.	25.0	0.2

In the Software Engineering Management KA, 46 studies and 19 research topics were found. This KA has 7.6% of the studies found in the SLR. In Table 10, an interesting observation is that the Software Engineering Management KA is covered mainly in the Models and Processes improvement topics.

Table 10: Topics of the Software Engineering Management KA.

Topics	Area%	Total%
Models; Process improvement;	13.0	1.0
Teams and organizations management; Challenges;	10.9	0.8
Tools and techniques; Development strategies;	8.7	0.7
Management and risk reduction; Approaches to projects with similar environments;	6.5	0.5
Business models;	4.3	0.3
Project Scope Management; Agile methodologies; Estimation of effort; Metrics; Ontologies; Middlewares; Support agents; Frameworks; Valuation in virtual environments; Security.	2.2	0.2

The Software Engineering Processes KA had the largest percentage of studies found in SLR, among SWEBOK KAs. This KA has 14.4% of the found studies (Table 11). In the Methods and Tools of

Table 11: Topics of the Software Engineering Processes KA.

Topics	Areas%	Total%
Management and improvement of processes;	26.4	3.8
Process and capacity models;	18.4	2.7
Teams and organizations management;	13.8	2.0
Tools;	5.7	0.8
Software development environments; Concepts of the GSE;	4.6	0.7
Methods; Agile methodologies; Process support systems; Knowledge management;	3.4	0.5
Collaborative Development; Metrics; Emerging topics and challenges;	2.3	0.3
Patterns; Frameworks; Software Quality; Components; UML diagrams; Software life cycles; Risks; User profiles; Performance analysis; Applications; Allocation of tasks.	1.1	0.2

Software Engineering KA, 69 studies were found, mapped into 24 topics. This KA has 11.4% of the studies found in the SLR (Table 12).

In the Software Quality KA 15 studies and 10 research topics were found. This KA has 2.4% of studies found in the SLR. The topics are divided in 3 groups, as presented in Table 13.

Table 12: Topics of the Methods and Tools of Software Engineering KA.

Topics	Areas%	Total%
Process support tools;	18.8	2.2
Communication tools;	11.6	1.3
Research on existing GSE tools;	10.1	1.2
Modeling of processes; Personnel management tools;	7.2	0.8
Project management tools;	5.8	0.7
Tools for requirements; Tools for real-time simulation;	4.3	0.5
Methods and tools for the Web; Tools for simulating projects; Time optimization tools; Tools for Assessment of case tools documentation; Tools for replacement and acquisition of the code;	2.9	0.3
Artificial intelligence; Tools for the development of teams skills; Construction of middleware; Quality tools; New methods unifying methodologies; Wiki; Route calculation tools; Testing tools; Offshore development tools.	1.4	0.2

Table 13: Topics of the Software Quality KA.

Topics	Area%	Total%
Projects quality management;	73.0	1.8
Quality assessments; Management and quality process; Influence of distance and geographical dispersion in software quality; Analysis of validation experiences;	13.3	0.3
Open source Software Quality; Frameworks; New approaches to Software Quality; Models; Performance indicators.	6.7	0.2

3.2.2 Research Topics of the PMBOK KAs

In Table 14, we present the main topics discussed in the Project Integration Management KA.

Table 14: Topics of the Project Integration Management KA.

Topics	Area%	Total%
Coordination models; Approaches;	22.2	0.3
Coordination processes; Life cycles; Knowledge management; Project phases; Offshore development tools.	11.1	0.2

In the Project Scope Management KA, 4 studies were identified, and each one has 1 topic associated. In Table 15, the topics are presented.

Table 15: Topics of the Project Scope Management KA.

Topics	Area%	Total%
Development routes studies; Group projects management; Models; Collaborative software.	25.0	0.2

In the Project Cost Management KA, 3 studies and only 1 topic were found, as can be seen in the Table 16.

Table 16: Topics of the Project Cost Management KA.

Topics	Areas%	Total%
Cost Analysis of collaboration models.	100.0	0.4

In the Project Quality Management KA, 11 studies were found, mapped into 7 research topics. This KA had 1.4% of the studies found in this SLR, and most studies discuss defect detection and quality evaluation (Table 17).

Table 17: Topics of the Project Quality Management KA.

Topics	Area%	Total%
Defect detection and quality evaluation;	27.3	0.5
Quality frameworks; Quality approaches;	18.2	0.3
Quality in offshore projects; Distance and software quality; Analysis of validation experiences; Quality and geographic dispersion.	9.1	0.2

In the Project Human Resource Management KA, 40 studies were mapped into 10 topics. This KA had 6.6% of the studies. In Table 18, the identified topics are presented.

Table 18: Topics of the Project Human Resource Management KA.

Topics	Area%	Total%
Teams management and development organizations;	37.5	2.5
Virtual teams;	20.0	1.3
Tools; Ability of teams and individuals;	12.5	0.8
Impact of the culture on the design team;	10.0	0.7
Models;	7.5	0.5
Adaptations of roles;	5.0	0.3
Allocation of tasks and teams; Simulation approaches; Communication practices.	2.5	0.2

In the Project Communications Management KA, 51 studies were found, corresponding to 8.4% of the studies. In this KA, 17 topics were identified, which are presented in Table 19.

Table 19: Topics of the Project Communications Management KA.

Topics	Area%	Total%
Communication support tools;	17.6	1.5
Communication flaws, barriers, case studies and experiments;	15.7	1.3
Communication management and coordination of knowledge; Communication patterns between project groups;	7.8	0.7
Communications requirements; Communication models; Agile methodologies;	5.9	0.5
Coordination mechanisms of communication; Structures of delay; Effects of distance in communication; Social networks and wiki;	3.9	0.3
Studies of communication language and linguistic differences; Use of instant messages in GSE; Predictive modeling processes using notations to improve communication among teams; Transcripts of communication among teams; Ontology for communication.	2.0	0.2

In the Project Risk Management KA, 11 studies were found and mapped into 10 topics. This KA had 1.8% of the studies found in this SLR. The identified topics in this KA are presented in Table 20.

Table 20: Topics of the Project Risk Management KA.

Topics	Area%	Total%
Analysis and coordination methods;	45.5	0.8
Frameworks;	36.4	0.7
Risk reduction; Risk assessment types; Identification of problems related to the emergence of risks; Risk management models; Agile methodologies; Risk Factors in outsourced offshoring;	18.2	0.3
Web Risks; Tools.	9.1	0.2

3.2.3 Research Topics in New KAs

In Table 21, we present the topics classified in the General Area:

Table 21: Topics of the General Area.

Topics	Area%	Total%
Education;	31.3	5.1
Revisions of workshops;	18.2	3.0
Challenges related to spatial and temporal aspects, cultural partners;	17.2	2.8
Lessons learned and practices;	12.1	2.0
New research areas;	9.1	1.5
Evolution of GSE;	4.0	0.7
Agile methodologies; Strategies for GSE;	3.0	0.5
Analysis of the technical and non-technical challenges for GSE; Difficulties for the research in the area;	2.0	0.3
Presentation of the research groups; Concepts; Benefits of using GSE.	1.0	0.2

In the Position Paper KA, 5 studies and 2 topics were identified (Table 22).

Table 22: Topics of the Position Paper KA.

Topics	Area%	Total%
Study prospects in GSE;	60.0	0.5
Considerations about culture, location, coordination in GSE.	40.0	0.3

In Table 23, the topics of the Systematic Review KA are presented. In this area, we found 8 studies mapped into 8 topics.

Table 23: Topics of the Systematic Review KA.

Topics	Area%	Total%
Challenges and solutions in GSE; Critical barriers in GSE;	25.0	0.3
Process models; Tools; Development patterns and practices; Aspects of communication; Success factors; Agile methodologies.	12.5	0.2

4 DISCUSSION

The studies that discuss GSE are focused mainly on the following KAs: Software Engineering Processes (14%), Methods and Tools of SE (11%), Project Communications Management (8%) and Project Human Resource Management (6%). The high percentage of studies in these KAs was not a surprise, because GSE to be successful is very dependent on these KAs.

The Project Procurement Management and Project Cost Management KAs, both from PMBOK, did not have identified studies. Other KAs, such as Time Management had less than 10 studies each.

The Software Engineering Processes, Methods and Tools of Software Engineering and Design software KAs from SWEBOOK correspond together to 27.1% of studies in GSE. These areas obtained a higher percentage of studies, in comparison to KAs of the PMBOK. This result indicates that most challenges of GSE might be concentrated in these KAs.

Based on this SLR we found studies for most Software Engineering KAs, but still there are topics not explored, and this might be an opportunity. We also identified new topics, classified in the General Area KA. These new topics indicate that there are several research opportunities in the GSE field, with several interesting topics to be explored, as the area evolve.

5 CONCLUSIONS

GSE is increasing and it is essential for organizations that wish to be in the software development market. Several organizations are doing GSE, and research in this field is essential to the growth of GSE as an area of research.

The existing studies are distributed over almost all SE KAs. Most of the studies found are concentrated in KAs that are the basis of GSE, such as Software Engineering Processes, Methods and

Tools, Project Communications Management and Project Human Resource Management.

The analysis of KAs' study topics reveals that the GSE is an immature research field and that there are still many research opportunities. This way, we believe that the development of this work contributed to the identification of emerging research areas and topics, along with research opportunities. In order to continue this work, this SLR may be expanded to reach other digital libraries. New studies can help identifying new research topics and additional study areas.

REFERENCES

- Abran A., Moore J. W. et al., 2004. *Guide to the Software Engineering Body of Knowledge (SWEBOOK®)*, IEEE Computer Society 2004 Guide.
- Ebling T., Audy, J. L. N., Prikladnicki, R., 2009. A Systematic Literature Review of Requirements Engineering in Distributed Software Development Environments. *ICEIS (3) 2009*, 363-366.
- Jiménez, M., Piattini, M., Vizcaino, A., 2009. Challenges and improvements in distributed software development: A systematic review. *Advances in Software Engineering*.
- Kitchenham, B., 2004. Procedures for Performing Systematic Reviews. Joint Technical Report SE0401 and NICTA Technical Report 0400011T.1, Software Engineering Group, Department of Computer Science, Keele University.
- PMBOK, A Guide to the Project Management Body of Knowledge (PMBOK® Guide), Third Edition, 2004, *Project Management Institute Publications*, U.S.A.
- Prikladnicki, R., Audy, J. L. N., Shull, F., 2010. Patterns in Effective Distributed Software Development. *IEEE Software*, v. 27, 12-15.
- Smite, D., Wohlin, C., Gorschek, T., Feldt R., 2009. Empirical evidence in global software engineering: a systematic review. *Empirical Softw. Eng.*, 91-118.
- Smite, D., Wohlin, C., Gorschek, T., Feldt, R., 2010. Empirical evidence in global software engineering: a systematic review. *Empirical Softw. Eng.* 2010, 91-118.