A Software Cost Estimation Taxonomy for Global Software Development Projects

Manal El Bajta and Ali Idri
Software Project Management Research Team, ENSIAS, Mohammed V University of Rabat, Morocco

Keywords: Software Cost Estimation, Global Software Development, Taxonomy.

Abstract: Nowadays, software cost estimation plays an important role in the management and development of distributed projects. The state of the art and cost estimation practice for Global Software Development (GSD) have recently been identified. This knowledge has still not been structured. The objective of this paper is to structure the knowledge about cost estimation for GSD. We used a design method to organize the knowledge identified as a cost estimation taxonomy for GSD. The proposed taxonomy offers a classification scheme for the cost estimation of distributed projects. The cost estimation taxonomy consists of four dimensions: cost estimation context, estimation technique, cost estimate and cost estimators. Each dimension in turn has multiple facets. The taxonomy could then be used as a tool to developing a repository for cost estimation knowledge.

1 INTRODUCTION

Software development effort estimation for Global Software Development (GSD) concerns the prediction of the effort needed to develop a global software project (Peixoto et al., 2010). Development effort is considered to be one of the major components of software costs, particularly as regards global development, and is usually the most challenging to predict (Lamersdorf et al., 2010), (Idri and Amazal, 2012).

To get a more comprehensive understanding of how software cost estimation is practiced by distributed teams, a follow-up study is carried out to elicit the state of the practice on cost estimation in GSD projects (El Bajta et al., 2015), (El Bajta et al., 2018). This study identified and aggregated knowledge on cost estimation in GSD from the literature and practice by means of a systematic review and a survey respectively. The knowledge includes aspects such as approaches used to estimate cost in GSD using size measures, cost drivers and the context in which estimates are carried out. This body of knowledge on cost estimation in GSD was structured in order to launch both future research and improve practice in this field. A taxonomy is a classification scheme used in Software Engineering (SE) to organize the body of knowledge (Glass and Vessey, 1995).

The main purpose of this paper is to design a software cost estimation taxonomy for GSD projects that allows managers to accurately assess proposed changes. Figure 1 illustrates the cost estimation process to build the cost estimate taxonomy. It consists on three steps: define cost drivers to be included in the study, collect the extracted data and knowledge, and build software cost estimation taxonomy.

This paper is structured as follows: we present related work in Section 2, identify the research methodology in Section 3, and report the results we obtained in Section 4. We identify threats to validity in Section 5 and present conclusions and thoughts on the future work in Section 6.

Figure 1: Cost estimation process.

2 RELATED WORK

Gumm (Gumm, 2006) developed a taxonomy of distribution to classify GSD projects dimensions. The taxonomy was used to provide a basis for discussing the challenges related to GSD projects and was based on an earlier study performed by the same author (Gumm, 2005). The proposed taxonomy uses four different dimensions (physical distribution, organizational distribution, temporal distribution and distribution among stakeholder groups measured on a high-medium- low scale) to classify the distribution of peo-

El Bajta, M. and Idri, A.
A Software Cost Estimation Taxonomy for Global Software Development Projects.
DOI: 10.5220/0007841202180225
In *Proceedings of the 14th International Conference on Software Technologies (ICSOFT 2019)*, pages 218-225
ISBN: 978-989-758-379-7

ple and artifacts in GSD context.

Laurent et al. (Laurent et al., 2010) proposed a taxonomy and a visual notation to address globally distributed requirements engineering projects. The main goal of the authors was to design the modeling language, including site locations, stakeholder roles, communication flows, critical documents, and supporting tools and repositories. The proposal was based on the findings of eight in-depth interviews with requirements analysts who worked on requirements elicitation, analysis, and specification tasks in globally distributed projects. These Interviews were performed with the team leaders responsible for eliciting and gathering the requirements in each project.

Smite et al. (Šmite et al., 2010) carried out a study specifically focused on evidence of empirical global software engineering. The study proposed a classification scheme to extract data from empirical studies and systematize existing empirical global software engineering studies. The proposed classification scheme helped to categorize the data extracted from the study population, empirical background and results. The study revealed that the collection of papers with empirical data on the subject of the desired systematic review was one of the main challenges, since globally distributed project is at the cutting edge of cross-disciplinary research. The process of deducting and collecting information about empirical background work was another difficulty reported. These highlight the need for thorough descriptions of contexts in which empirical studies are conducted.

More recently, Smite et al. (Šmite et al., 2014) have proposed a taxonomy of sourcing procedures. The taxonomy proposed gives a typical terminology and takes into consideration the classification of GSD projects with spotlight on the sourcing strategies (e.g., Offshore outsourcing, offshore insourcing). The result provided a systematically accumulated set of terms categorized in the form of a taxonomy.

The taxonomy proposed by Smite et al. (Šmite et al., 2014) is considered as knowledge classification approach. These taxonomy of sourcing strategies is the most closely related work to ours, hence our decision to use these taxonomy to understand the GSD project setting. The base of these taxonomy is more exhaustive, giving a more extensive scope of relevant dimensions and clear criteria to classify GSD projects. Furthermore, this taxonomy was also developed with the participation of several GSD experts, which gives the taxonomy more credibility and validity. Therefore, our goal is to propose a taxonomy to classify all dimensions of cost estimation for GSD, and include categories related to empirical focus, subjects of investigation and sources of data collection.

3 RESEARCH METHODOLOGY

This section describes the research question and methodology used to design and evaluate the proposed taxonomy.

3.1 Research Questions

This paper addresses one Research Question (RQ), which is presented below:

RQ: How to organize the knowledge on cost estimation for GSD projects?

The RQ is answered by organizing the cost estimation knowledge for GSD projects as a taxonomy.

3.2 Taxonomy Design

The focus of this subsection is to present the method to design the software cost estimation taxonomy used in this study. Usman et al. (Usman et al., 2017) present a revised and updated method on taxonomies in the field of software engineering. As shown in Table 1, the method consists of four phases and thirteen activities .

Phase 1: Planning

Planning represents the first phase wherein basic decisions about the taxonomy implementation and design are made. In this phase, six activities are defined as shown in Table 1.

In activity **A1**, the SE knowledge area is selected and described to make easier the understanding of the context of the taxonomy and thus its application. The taxonomy proposed in this paper is about cost estimation for GSD context. During release and planning phases, cost estimation plays an important role in the management of distributed projects. Cost estimation falls inside the extent of the "Software engineering management" knowledge area in SWEBOK version 3 (Bourque et al., 2014).

In activity **A2**, the main objectives and scope of the taxonomy are to propose a classification scheme that can be used to describe software cost estimation activities for GSD projects. A number of studies, included in the Systematic Mapping Study (SMS) on software cost estimation for GSD (El Bajta et al., 2015) did not report significant information on the main context, techniques and also predictors used during cost estimation. Researchers and practitioners could therefore use the proposed taxonomy to consistently report and recall important aspects related to software cost estimation for GSD projects.

Table 1: Taxonomy design method.

Phase	Activity
Planning	A1. Define SE knowledge area A2. Describe the objectives of the taxonomy A3. Describe the subject matter to be classified A4. Select classification structure type A5. Select classification procedure type A6. Identify the sources of information
Identification and extraction	A7. Extract all terms A8. Perform terminology control
Design and construction	A9. Identify and describe taxonomy dimensions A10. Identify and describe categories of each dimension A11. Identify and describe the relationships A12. Define the guidelines for using and updating the taxonomy
Validation	A13. Validate the taxonomy

In activity **A3**, the subject matter for the classification defines what exactly is classified in the taxonomy. The cost estimation activities of the globally developed projects is the subject of this taxonomy.

In activity **A4**, an appropriate classification structure type is selected. Four basic classification structures are defined: hierarchy, tree, paradigm and faceted analysis (Kwasnik, 1999). To structure our taxonomy, faceted classification is selected, since it is suitable for evolving areas, such as software cost estimation for GSD. The subject matter is classified from several perspectives (facets) in faced classification-based taxonomies. Each facet of the proposed facets is considered independent and has its own attributes, making it easy to develop facet-based taxonomies (Kwasnik, 1999).

In activity **A5**, an appropriate classification procedure is determined. These type can be qualitative, quantitative or both. Every facet of our taxonomy has a set of values. The qualitative procedure is used to select relevant facet values on the basis of extracted data to characterize a specific estimation activity. In some cases, it is impossible to assign a value simply because of insufficient data.

In activity **A6**, the data sources and data collection methods are identified to facilitate the prospection of knowledge related to the subject matter and taxonomy. These data sources are selected from peer-reviewed empirical studies on cost estimation for GSD published in literature.

Phase 2: Identification and Extraction

In this phase the relevant data required by the organization is identified and extracted. Two activities are defined as shown in Table 1.

In activity **A7**, the terms and concepts relevant to the taxonomy are extracted from the sources identified in the first phase “planning”.

In activity **A8**, inconsistencies in the extracted data are identified and removed.

Phase 3: Design and Construction

In this phase, the taxonomy is designed and constructed in order to identify dimensions and categories in which the data items extracted can be organized. Four activities are defined as shown in Table 1.

In activity **A9**, the taxonomy dimensions are identified and described. They represent the main dimensions or perspectives under which subject matter entities are classified. Facet-based classification taxonomy structure must have multiple dimensions (at least two dimensions).

In activity **A10**, the categories for each of the dimensions are identified and described, each dimension must have at least two categories.

In activity **A11**, the relationships between categories and dimensions shall be identified and described. Note that in some cases there are no relationship between dimensions, i.e. this activity could be skipped.

In activity **A12**, the guidelines are provided to facilitate the adoption and evolution of the taxonomy.

Phase 4: Validation

Validation represents the last phase of this method to ensure that the designed taxonomy is useful for users to achieve their goals. In this phase, only one activity is defined as shown in Table 1.

In activity **A13**, the taxonomy can be validated through benchmarking. Since there is no existing taxonomy on software cost estimation for GSD, benchmarking our taxonomy against existing ones is impossible.

4 RESULTS

In this section, we present the software cost estimation taxonomy to answer the **RQ:** Organizing the knowledge on cost estimation in GSD.

Taxonomies represent an effective tool to organize and communicate the knowledge in an area (Glass and Vessey, 1995). We have organized the identified knowledge on cost estimation for GSD as a taxonomy. The proposed taxonomy was developed according to the method presented in Section 3. In this subsection, we describe the results of this method, i.e. the software cost estimation taxonomy.

Four dimensions are extracted and placed at the top level of the taxonomy, as shown in Figure 2. A dimension consists of all facets that are interrelated. These four dimensions characterize the first level of taxonomy, and give an outline of the taxonomy at a more extensive level.

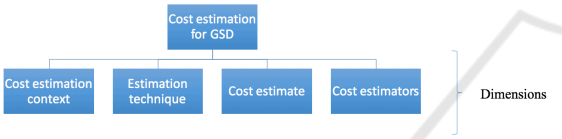

Figure 2: First level of cost estimation taxonomy for GSD.

The estimation context represents the collection of those facets that define and characterize the context in which the estimation activity in a distributed project is carried out. So as to fully characterize a specific software cost estimation activity of GSD project, facets of all dimensions should to be described. Each of these four dimensions and their facets are described in detail below.

4.1 Cost Estimation Context

The context is a central concept in empirical software engineering. It is one of the distinctive features of the discipline and it is an inseparable part of software practice. Context refers to a broad perspective, and it needs to be properly captured, reported , and contextualized in the empirical SE studies to communicate the applicability of the research findings. Thus, context draws attention to what resources are nearby, and when and where to use the reported findings (Dybå et al., 2012), (Dey, 2001).

Seven facets are extracted from context dimension. These facets and their possible values are presented in Figure 3. They are described below:

- *Planning:* Estimation supports planning at various levels in GSD context. This mainly includes release and sprint planning (Hossain et al., 2009), while some teams may also make estimates during daily meetings. Project bidding is another level at which companies must estimate the total development cost in advance to bid for the projects.

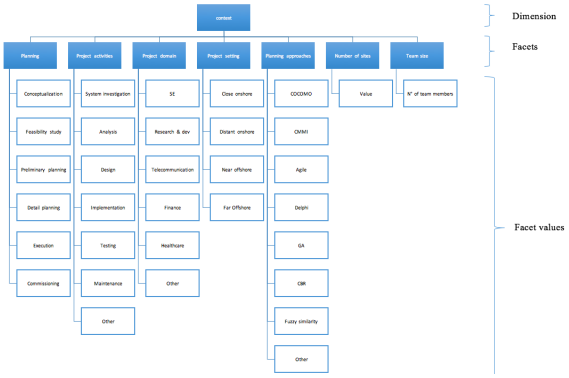

Figure 3: Cost estimation context dimension.

- *Project Activities:* This facet shows which development activities are accounted for in the estimation of software costs. For example, The product life cycle describes maintenance activity, or the total cost estimate do not include the time spent on maintenance.
- *Project Domain:* This facet captures the domain of project for which the software cost development is being estimated. Different domains could lead to different sets of cost estimation. We have used the categories from the project domains reviewed in previous study (El Bajta et al., 2017b).
- *Project Setting:* This facet represents the setting in which the global software teams are developing the project. Smite et al. (Šmite et al., 2014) proposed a global software engineering taxonomy that characterizes two broad settings of the global teams: onshore and offshore.
- *Planning Approaches:* This facet documents the planning approach practiced by GSD team.
- *Number of Sites:* It records the number of sites for the GSD project, and thus communicates important information related to the different sites worldwide.
- *Team Size:* It documents the team size which is responsible for developing the estimated tasks.

4.2 Estimation Technique

This dimension includes the facets that are related to estimation techniques. Those facets should be reported to characterize a GSD team's estimation activity (El Bajta, 2015). Figure 4 describes the facets of this dimension and their corresponding values.

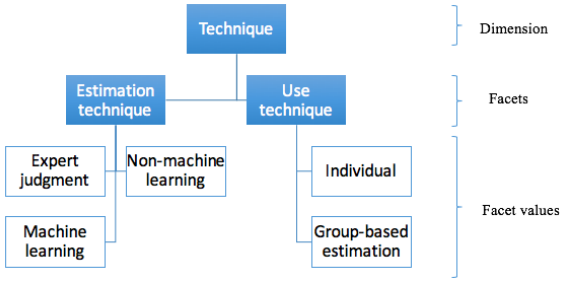

Figure 4: Estimation technique dimension.

- *Estimation Technique:* This documents the estimation techniques applied for GSD projects. According to the SMS results (El Bajta et al., 2015), the cost estimation techniques for GSD projects are expert judgment, machine learning and non-machine learning (Hosni et al., 2017), (El Bajta et al., 2017a), (Amazal et al., 2014b), (Idri et al., 2016d), (Idri et al., 2016c).
- *Use Technique:* In GSD, there are different types of cost estimation techniques. An individual or a group of experts can use these techniques. This facet documents whether the effort was estimated using an estimation technique based on individuals or groups (Amazal et al., 2014a), (Idri et al., 2002b).

4.3 Cost Estimate

The main output of the estimation activity is cost estimates. The facets proposed in this dimension define cost estimate. Those facets with their corresponding values are presented in Figure 5.

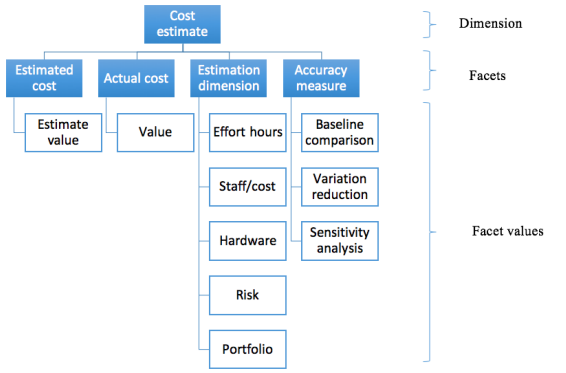

Figure 5: Cost estimate dimension.

- *Estimated Cost:* This facet documents the estimated cost that represents the main output of the estimation activity.
- *Actual Cost:* It is important to have the actual cost at the end of planning, to enable comparison with the estimated cost.
- *Estimation Dimension:* This facet documents the important and critical dimensions of estimation, e.g. estimation of development effort as total effort hours.
- *Accuracy Measure:* This facet records cost performance ways to assess the accuracy of the applied estimation technique (Idri et al., 2016b).

4.4 Cost Estimators

Cost estimators play an important role in calculating costs. They consist of cost drivers such as size, team capabilities, product requirements, etc (Idri et al., 2016a), (Idri and Abran, 2001). One of the most required development cost is related to project size. Five facets are collected regarding the cost estimators dimension. These facets and their possible values are presented in Figure 6.

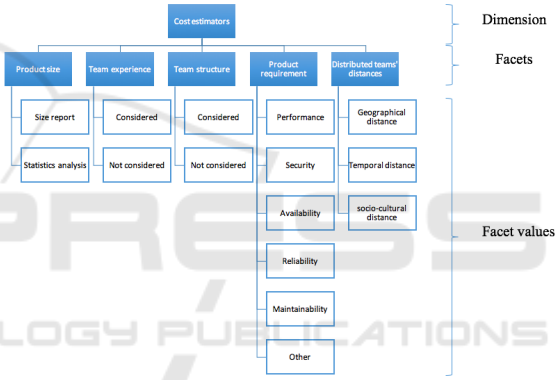

Figure 6: Cost estimators dimension.

- *Product Size:* In general, the development cost is strongly correlated with product size (Jaakkola, 2009). This facet documents whether distributed teams use size as an estimator and which statistical analysis is used to represent this size.
- *Team Experience:* A development team experiences with global software development projects impact the required cost (Ramasubbu et al., 2011). This facet describes whether a team experience was considered or not in arriving at the cost estimates.
- *Team Structure:* Distribution of skills and team structure impact the required effort (Agerfalk et al., 2005). This facet documents whether the structure of the team members was considered or not during the cost estimation session.
- *Product Requirement:* Strict product requirements increase the development cost (Chung et al., 2012), (Idri et al., 2002a). This facet records

which product requirements were considered in arriving at cost estimates.

- *Distributed Teams' Distances:* The geographical, temporal, and socio-cultural distances between global development teams increase the development cost due to the increased complexity of the collaboration and communication (Holmstrom et al., 2006).

5 THREATS TO VALIDITY

The results of the proposed taxonomy may have been influenced by the coverage of the study search, and also the inaccuracy in study data extraction. Four types of validity threats (Easterbrook et al., 2008) of the study results are therefore discussed in the following subsections.

Construct validity is concerned with issues caused by poor data collection and recording, also, exactitude of the interpretation of the concepts studied and the completeness of the relevant studies collected. In this study, the data identified and aggregated are extracted from the SMS (El Bajta et al., 2015) and survey (El Bajta et al., 2017b). The extracted data was used as the main input for the taxonomy designed in this study. To ensure the correct interpretation of these data, we checked the contributions of the concepts in related literature and all the authors discussed these data in order to reach a consensus as to their use and contribution.

Internal validity is concerned with analyzing extracted data. Threats to internal validity are important as our study does not present an evaluation of the taxonomy. In future studies, we plan to evaluate and assess the taxonomy by using it to characterize estimation cases from the literature and the industry.

Conclusion validity is to ensure that reasonable conclusions are drawn on the basis of data collected and that problems such as the bias of researchers do not lead to incorrect conclusions. We used a taxonomy design method to systematically organize the knowledge on software cost estimation for GSD as taxonomy in a systematic manner.

External validity is concerned with the representativeness of the selected studies as regard the overall goal of the study. The results of this study were considered with regard to the cost estimation for GSD projects. These results can serve as a starting point for researchers and practitioners working in this field to further improve the completeness and usefulness of the proposed taxonomy.

6 CONCLUSIONS

The development of taxonomies helps to structure, generalize and share existing knowledge and to advance research (Glass and Vessey, 1995). We have organized the existing body of knowledge on cost estimation in GSD as a taxonomy. The taxonomy has been systematically developed by following a taxonomy design method. One research question was addressed by incorporating five dimensions to organize knowledge on cost estimation for GSD projects. The main usage for our taxonomy is to provide a basis for researchers to classify their own studies and related studies on cost estimation for GSD field. This taxonomy could therefore be used as a tool to develop a cost estimation knowledge repository to better understand and improve the cost estimation practice in the global development context in the long term.

The usefulness of the taxonomy has not been demonstrated in the study. We plan to apply the developed taxonomy on data extracted and reported in the literature to characterize cost estimation cases.

REFERENCES

Agerfalk, P. J., Fitzgerald, B., Holmstrom Olsson, H., Lings, B., Lundell, B., and Ó Conchúir, E. (2005). A framework for considering opportunities and threats in distributed software development.

Amazal, F.-A., Idri, A., and Abran, A. (2014a). An analogy-based approach to estimation of software development effort using categorical data. In *2014 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement*, pages 252–262. IEEE.

Amazal, F. A., Idri, A., and Abran, A. (2014b). Software development effort estimation using classical and fuzzy analogy: a cross-validation comparative study. *International Journal of Computational Intelligence and Applications*, 13(03):1450013.

Bourque, P., Fairley, R. E., et al. (2014). *Guide to the software engineering body of knowledge (SWEBOK (R)): Version 3.0*. IEEE Computer Society Press.

Chung, L., Nixon, B. A., Yu, E., and Mylopoulos, J. (2012). *Non-functional requirements in software engineering*, volume 5. Springer Science & Business Media.

Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, 5(1):4–7.

Dybå, T., Sjøberg, D. I., and Cruzes, D. S. (2012). What works for whom, where, when, and why?: on the role of context in empirical software engineering. In *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement*, pages 19–28. ACM.

Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering*, pages 285–311.

El Bajta, M. (2015). Analogy-based software development effort estimation in global software development. In *IEEE 10th International Conference on Global Software Engineering Workshops (ICGSEW)*, pages 51–54. IEEE.

El Bajta, M., Idri, A., Fernández-Alemán, J. L., Ros, J. N., and Toval, A. (2015). Software cost estimation for global software development a systematic map and review study. In *Evaluation of Novel Approaches to Software Engineering (ENASE), 2015 International Conference on*, pages 197–206. IEEE.

El Bajta, M., Idri, A., Ros, J. N., Fernández-Alemán, J. L., De Gea, J. M. C., Garcıa, F., and Toval, A. (2018). Software project management approaches for global software development: a systematic mapping study. *Tsinghua Science and Technology*, 23(6):690–714.

El Bajta, M., Idri, A., Ros, J. N., Fernandez-Aleman, J. L., and Toval, A. (2017a). Estimation of costs and time for the development of distributed software. In *Software Project Management for Distributed Computing*, pages 25–42. Springer.

El Bajta, M., Idri, A., Ros, J. N., Fernández-Alemán, J. L., and Toval, A. (2017b). Software cost attributes in global software development projects: A survey. In *Proceedings of the 9th International Conference on Information Management and Engineering*, pages 96–101. ACM.

Glass, R. L. and Vessey, I. (1995). Contemporary application-domain taxonomies. *IEEE Software*, (4):63–76.

Gumm, D. C. (2005). The phenomenon of distribution in software development projects: A taxonomy proposal. In *Proc. European Mediterranean Conf. Information Systems*. Citeseer.

Gumm, D. C. (2006). Distribution dimensions in software development projects: A taxonomy. *IEEE software*, 23(5):45–51.

Holmstrom, H., Conchúir, E. Ó., Agerfalk, J., and Fitzgerald, B. (2006). Global software development challenges: A case study on temporal, geographical and socio-cultural distance. In *Global Software Engineering, 2006. ICGSE'06. International Conference on*, pages 3–11. IEEE.

Hosni, M., Idri, A., and Abran, A. (2017). Investigating heterogeneous ensembles with filter feature selection for software effort estimation. In *Proceedings of the 27th International Workshop on Software Measurement and 12th International Conference on Software Process and Product Measurement*, pages 207–220. ACM.

Hossain, E., Babar, M. A., and Paik, H.-y. (2009). Using scrum in global software development: a systematic literature review. In *Global Software Engineering (ICGSE), 2009 Fourth IEEE International Conference*, pages 175–184. IEEE Computer Society.

Idri, A., Abnane, I., and Abran, A. (2016a). Missing data techniques in analogy-based software development effort estimation. *Journal of Systems and Software*, 117:595–611.

Idri, A. and Abran, A. (2001). A fuzzy logic based set of measures for software project similarity: validation and possible improvements. In *Proceedings Seventh International Software Metrics Symposium*, pages 85–96. IEEE.

Idri, A., Abran, A., and Khoshgoftaar, T. M. (2002a). Estimating software project effort by analogy based on linguistic values. In *Proceedings Eighth IEEE Symposium on Software Metrics*, pages 21–30. IEEE.

Idri, A. and Amazal, F. A. (2012). Software cost estimation by fuzzy analogy for isbsg repository. *Uncertainty Modeling in Knowledge Engineering and Decision Making*, pages 863–868.

Idri, A., Amazal, F. A., and Abran, A. (2016b). Accuracy comparison of analogy-based software development effort estimation techniques. *International Journal of Intelligent Systems*, 31(2):128–152.

Idri, A., Hosni, M., and Abran, A. (2016c). Improved estimation of software development effort using classical and fuzzy analogy ensembles. *Applied Soft Computing*, 49:990–1019.

Idri, A., Hosni, M., and Abran, A. (2016d). Systematic literature review of ensemble effort estimation. *Journal of Systems and Software*, 118:151–175.

Idri, A., Khoshgoftaar, T. M., and Abran, A. (2002b). Investigating soft computing in case-based reasoning for software cost estimation. *International Journal of Engineering Intelligent Systems for Electrical Engineering and Communications*, 10(3):147–157.

Jaakkola, H. (2009). Towards a globalized software industry. *Acta Polytechnica Hungarica*, 6(5):69–84.

Kwasnik, B. H. (1999). The role of classification in knowledge representation and discovery.

Lamersdorf, A., Munch, J., Fernandez-del Viso Torre, A., Sanchez, C. R., and Rombach, D. (2010). Estimating the effort overhead in global software development. In *Global Software Engineering (ICGSE), 2010 5th IEEE International Conference on*, pages 267–276. IEEE.

Laurent, P., Mader, P., Cleland-Huang, J., and Steele, A. (2010). A taxonomy and visual notation for modeling globally distributed requirements engineering projects. In *2010 5th IEEE International Conference on Global Software Engineering*, pages 35–44. IEEE.

Peixoto, Lima, C. E., Audy, J. L. N., and Prikladnicki, R. (2010). Effort estimation in global software development projects: Preliminary results from a survey. In *Global Software Engineering (ICGSE), 2010 5th IEEE International Conference*, pages 123–127. IEEE.

Ramasubbu, N., Cataldo, M., Balan, R. K., and Herbsleb, J. D. (2011). Configuring global software teams: a multi-company analysis of project productivity, quality, and profits. In *Software Engineering (ICSE), 2011 33rd International Conference on*, pages 261–270. IEEE.

Šmite, D., Wohlin, C., Galviņa, Z., and Prikladnicki, R. (2014). An empirically based terminology and taxonomy for global software engineering. *Empirical Software Engineering*, 19(1):105–153.

Šmite, D., Wohlin, C., Gorschek, T., and Feldt, R. (2010). Empirical evidence in global software engineering: a systematic review. *Empirical software engineering*, 15(1):91–118.

Usman, M., Britto, R., Börstler, J., and Mendes, E. (2017). Taxonomies in software engineering: A systematic mapping study and a revised taxonomy development method. *Information and Software Technology*, 85:43–59.