Knowledge and Contextual Information Management in Global Software Development: Challenges and Perspectives

Elisa Hatsue Moriya Huzita¹, Gislaine Camila Lapasini Leal², Renato Balancieri¹, Tânia Fátima Calvi Tait¹, Edwin Cardoza², Raqueline Ritter de Moura Penteado¹, Rafael Leonardo Vivian¹

Department of Informatics, State University of Maringá, Brazil¹
Department of Production Engineering, State University of Maringá, Brazil²

¹{emhuzita, renao, tait, raque}@din.uem.br¹
rafavivian@gmail.com¹

²{gclleal, evcgaldamez}@uem.br²

Abstract — Global Software development approach has become a reality with growing demand for complex software. However, this approach requires a greater attention because it demands additional coordination and communication among the distributed team members. So, tools and techniques to share information and knowledge generated by each team geographically dispersed are necessary. In this scenario, this position paper presents a set of challenges and opportunities for research in global software development (GSD). These challenges and opportunities are focused on contextual information management, knowledge management, performance management, socio-cultural aspects of GSD, and strategies to store and sharing data. They were mapped into a proposed research agenda with the following items: knowledge management in GSD projects; definition of mechanisms to support the capture of contextual information; use of cloud computing paradigm to store data generated; study of techniques for dealing with the contextual information; and establishment of performance dimensions to provide support for the decision maker that manages the distributed teams. The research agenda also includes strategies defined aiming to strengthen the partnership between university and industry. We believe that the collaboration between university and industry can contribute for the consolidation of the area in Brazil.

Keywords— contextual information, knowledge management, performance management, distributed software development, research agenda, university and industry collaboration

I. INTRODUCTION

Software development is characterized by increased complexity of projects and demand for reduced delivery time. These have corroborated for schedule and estimated cost of project go beyond that expected and thus do not meet the quality requirements as specified by the customer [1]. The growing need for making faster and more accurate to carry out several activities has encouraged people to collaboratively work together. However challenges regarding group work such as skills to cooperate and coordinate arise when activities are carried out involving more than one person. So, tools and practices to meet this demand are required [2]. The need for collaboration among teams geographically dispersed promoted the practice of global software development (GSD).

According to Carmel and Agarwal [3], Hyysalo et al [4] and Noll et al [5], GSD approach has been increasingly used. This approach has benefits regarding cost and quality since companies look for the best skilled professionals, which can be scattered around the world. Because of this distribution, the task of developing software becomes something almost uninterrupted. But, due to cultural, political and organizational differences, as well as temporal and geographical distances, challenges related to cooperation, coordination, and communication arise. So, in order to minimize these differences and distances and, yet make sure that individuals who are geographically distributed are collaborating, it is essential to have an infrastructure that ensures the exchange of information and knowledge among all engaged [6].

Due to physical distance in GSD, to obtain contributions from other team members, it is necessary to know not only the cooperation object, but also the way in which it was produced [7]. This information reflects the context, which addresses the circumstance under which the activities were developed. Contextual information can be exploited to detect cases of potential collaboration among several developers, which can then interact with each other. Therefore, it is important to emphasize that the absence of adequate mechanisms for sharing contextual information and also the knowledge generated at each site makes it difficult to exploit the benefits of global software development.

The captured information needs to be adequately represented, stored, processed and then disseminated for use in organization. Such information are related to business rules (domain), artifacts generated (documents), process (organizational, fundamental and support), technologies and tools, contextual information identified in software artifacts created or modified, as well as skills that a certain developer or
organization has to perform a specific activity in the software development process, among others. Furthermore it is necessary to store the decision and reasons of practices adopted by project manager and/or team for software development, improvement and maintenance [8]. So, the idea is to minimize the problems related to time and cost, improve software quality and offer adequate support for software project decision-making.

From the perspective of process integration of several organization/teams in GSD, it is also important to develop and to adopt practices of organizational performance management. It supports decision-making as well as contribute to improve the performance of critical dimensions (cost, time, quality) of GSD. Performance management is based on performance metrics, which results can influence team behavior aspects and promote continuous innovation of process and products [9]. In this scenario, performance management offers support for development, reuse and knowledge capitalization [10].

The GSD approach has become increasingly present in both academy and industry. In the academy, there are several research groups in Brazil, such as UEM, PUCRS, UFPE, UFPA, UFPE, among others, taking initiative, developing projects that could contribute towards the development of this area. The group of studies in distributed software engineering at the Universidade Estadual de Maringá (UEM) has been developing research in this area since 2002, with the goal of building an environment that supports GSD.

The works carried out over the years have explored several aspects of GSD, including: integrated approach for testing and distributed software development, artifacts version control, human resources allocation, risks management, project management, requirements in GSD, model of awareness in GSD environment, definition of a security politic for a GSD environment, among others. The scientific experience and the knowledge base generated by our group members that involves researchers, graduate and undergraduate students jointly with an adequate bibliographic research, make possible identify and motivate us to carry out research about knowledge management in Global Software Development.

In this context, this is a position paper expresses the opinion of our research group on prospects for research on GSD. The purpose of is to share a set of challenges and research opportunities identified by our group throughout the years considering the knowledge and contextual information management in GSD, and further discussion about the proposed research agenda. The remainder of the papers is organized as follow: Section II presents the theoretical background. Section III discusses the identified challenges and research opportunities. Section IV presents agenda for research opportunities. Section V concludes the paper presenting our final considerations.

II. BACKGROUND

Usually, software development organizations are characterized by the following aspects: (i) intellectual capital which constitutes one of the major assets; (ii) the recovery of organization member experience that can be used to increase productivity; (iii) the use of knowledge base for systematic resolution of problems; (iv) setting best practices to accelerate human resources training; and (v) the need to develop intellectual capital to suit the dynamicity of technologies evolution. It is also noted that in software industry is generated a considerable volume of technical data about the product and process adopted. So, it is necessary to establish mechanisms for managing knowledge and to promote the improvement of organizational performance [11].

According to Davenport and Prusak [12], knowledge can be described as a dynamic mix of experience, values, contextual information and expert insights. This knowledge provides a framework for evaluating and aggregate new experiences and information. Thus, the knowledge can be and is assessed by decisions or actions that resulted in this. For example, the reuse of knowledge enables to increase efficiency in the product development.

Human knowledge can be classified into explicit and tacit. Explicit knowledge can be externalized in formal language. Tacit knowledge corresponds to knowledge based on personal experience and involves intangible factors as beliefs, perspectives and values. The transfer of tacit knowledge to explicit can occur through one of the following knowledge conversion processes: socialization, externalization, combination and internalization [13].

Knowledge Management (KM) is a human activity that is part of the knowledge of an agent or a set of agents. It is a continuous network for interactions among agents aimed at handling, management, controlling, coordination, planning and organization of other agents, components and activities, through a process of generation and integration of knowledge. So, the goal is to produce a unified view that is well planned and directed, and thus, generating, maintaining, improving, acquiring, and broadcasting the knowledge base of the organization [11].

In GSD the software project manager needs to have new skills and concerns associated with those habitual in management area (plan, lead, coordinate and monitor) to share physical and human resources. Usually, project management involves technical and human factors. So, to deal with such factors it is necessary to know tools and techniques based on knowledge that provide support in planning and monitoring (performance and evaluation) of projects.

According to Kalling [10], knowledge is not usually used and even when it is used cannot be guaranteed the process improvement and products performance. So, it is essential to encourage the use of elements that convert knowledge into management mechanisms for improvement of activities. In this sense, performance management (metrics) can be proposed as a practice that adds value both individual and organizational knowledge.

Organizational Performance Management is a practice of continuous innovation that proposes the establishment, deployment, and use of performance metrics [14], [15], [16], [17]. It helps in the identification of critical activities and GSD process improvement. It also allows evaluate the results obtained on the cost, time, productivity, quality, customer satisfaction, and employees among other dimensions. In
addition, performance measurement will provide support to evaluate the effectiveness of practices related to knowledge management in software development organizations.

Several papers in the literature deal with knowledge management in GSD (e.g., [18], [19], [20], [21]). A hybrid approach for knowledge management in GSD is presented in Desouza et al [18]. Their study focuses on the reuse of global knowledge and on connection among local knowledge sources. This structure allows the use of knowledge and providing expertise index for each project, which makes it easier to recover knowledge. An architecture for knowledge management systems aimed at facilitating knowledge sharing and also to facilitate the recovery process using natural language queries and notifications is proposed in Shiva et al [19]. The authors emphasize that in GSD environments, where the knowledge can be stored in different formats and platforms, it is important to have a tool that can integrate, correlate and allow the knowledge extraction in an easy way. According to Liu et al [20] the socialization occurs in GSD when developers share their observations and personal work experience. Externalization occurs when these experiences are documented. Combination occurs when there is the coordination with other sectors of organizations. Internalization occurs when the documents are shared through interaction and with a process of trying and errors of different aspects of tacit knowledge. According to Dingsoy et al [21], identify practices and tools for knowledge management in GSD are more relevant due to missing face to face interaction, facilitate the work coordination and mutual learning among disperse teams.

The current literature on GSD presents some systematic reviews on several topics, such as: challenges and improvement in GSD [22], process models in the practice of GSD [23], and context-awareness on software artifacts in GSD [24] among others. Interesting concerns about GSD that could be explored in research are pointed by these authors.

III. CHALLENGES AND RESEARCH OPPORTUNITIES

In this section we present and discuss the challenges and opportunities of research that we have identified throughout the years in our research group. These elements are identified from the perspective of following aspects: contextual information management, knowledge management, performance management, social and cultural aspects involved in project management, and strategies to store and make available the data. We expect to collaborate with proposed solutions to deal with challenges related to communication, coordination, collaboration and performance in GSD.

Trends show that the size and complexity of software system will continue to grow up. This drives the search by resources and techniques that provide adequate support to development in distributed way [3], [4], [5]. Although this development approach brings some advantages, some challenges arise too. Among these challenges, communication is one of the most discussed ones. Communication allows the developers to share ideas, know the task status and resources, and also the activities of other team members. The context is a fundamental source of information that can be captured: by sensors, in software products, in human interactions, and others during the software development. Project and business decisions that have been adopted in several projects can be also considered. After the capture, the contextual information should be adequately represented, processed and used. So, the development of tools and mechanisms to deal with contextual information considering the scenario of global software development becomes a challenge.

The storage and use of the information over the years can make them in common practices adopted by people and organizations, charactering the knowledge generated. According to Nonaka and Takeuchi [13], the knowledge is highly personal and difficult to formalize, making it difficult to share. While the explicit knowledge is more related to informational process, the tacit has the intuitive nature and is deeply rooted in actions and experiences of an individual.

The contextual information can be in several situations that occur during the collaborative software development, such as in activities, artifacts or skills of participants. The context has the capacity to provide a major meaning to knowledge, promoting a more effective comprehension about a determined situation in the collaborative work [25]. Thus, the awareness about contextual information should be enlarged to increase the support into performing of activities in collaborative work.

In global software development environment, in addition to codifying, storing and sharing knowledge it is also necessary to provide awareness mechanisms that support participants to understand contextual information. The knowledge management should be considered together with the context of situation in which the knowledge was captured [25]. Thus, to understand a situation, it is necessary to know everything that is relevant to that condition.

Most of the problems about the availability of knowledge in organizations fall upon: (i) problems with knowledge transfer; (ii) errors due to lack of knowledge; (iii) critical knowledge in the hands of a few people; (iv) inability to perform the measurement in relation to knowledge use; (v) loss of relevant knowledge at critical time; and (vi) lack of knowledge sharing processes. These problems can become larger when GSD is adopted, since communication, collaboration and coordination are also part of the scenario in which the developers are.

Thus, the knowledge must be formalized, distributed, shared and applied across the organization. It can be converted into an asset that has value to the organization. Knowledge management can be used to promote the process improvement and thus ensure the quality of the product developed.

A major challenge to the efficient transfer of knowledge in GSD is due to the fact that team members are geographically distant, working with different time zone, with cultural difference and different experiences. So, the management and transfer of knowledge are critical success factors in this scenario. This management can be understood as a process to continually create new knowledge, disseminate it in the organization, incorporate it quickly on new products/services, technologies and systems [26]. The knowledge transfer consists in providing background information on software projects for people who need to know about these [27].
The concept of knowledge transfer is difficult to capture because there is no clear distinction between transferring knowledge and creating new knowledge [28]. The 'transfer' of knowledge can occur among people, organizations or countries.

Knowledge transfer is strategically important to organizations because it conveys the "know-how" which is located in a subunit where it was generated for other locations in organizations. Knowledge transfer also facilitates the workflow coordination by connecting multiple subunits that are geographically dispersed. Moreover, they can enable organizations to capitalize on business opportunities requiring the collaboration of several subunits. Finally, knowledge transfer is also crucial to the understanding among the various parts involved in software development. This knowledge can be synchronized and used to perform activities.

Therefore, the use of Knowledge Management in a business environment is important since the organization becomes the keeper of knowledge shared by its employees, making it more accessible and less dependent on people. Furthermore, the knowledge generated from previous software projects, which may be similar to the current project, may give a preview of the problems that may occur and also support in solving questions about the project more quickly.

In GSD, members of development teams can and should interact collaboratively to carry out their activities. It is in these interactions that context information, and also the knowledge that each one holds can be embedded [29], [30]. Therefore, regardless of the size of organizations involved, it is important to properly capture, represent, transfer, share, and finally to manage knowledge and contextual information.

Due to the large volume of data generated during the software development, especially in GSD, it is important to define the technologies that should be used to ensure reliability and availability of information. In this context infrastructure and storage services offered in cloud can be used for storing contextual information. The paradigm of cloud computing is a business model that offers on-demand services with high availability, reliability, elasticity and scalability. Therefore, users can access via network services anytime and anywhere by paying only the resources used. This implies zero upfront infrastructure investment, just-in-time infrastructure, more efficient resource utilization, usage-based costing and a real potential for shrinking the processing time [31]. One question that should be considered is the type of cloud: public, private or hybrid. The private provides more security because it is established in loco, but limits the benefits mentioned above because it depends on the IT infrastructure and employees of the company. According to Hashmi et al [32], the cloud paradigm can be used to facilitate efficient collaboration among geographically distributed teams during software development.

Another point to observe is the growing number of organizations that have adopted GSD. This scenario makes it suitable and appropriate for the academy to seek closer contact with the industry, establish partnerships and thus, together they can identify problems, analyze them and then propose feasible solutions grounded in scientific rigor but also useful for organizations involved in the GSD.

IV. AGENDA FOR RESEARCH OPPORTUNITIES

This section presents a research agenda from contextual information and knowledge management in global software development. This agenda is proposed based on the elements identified in the background section and in experience of our group in conducting GSD research.

- To identify, extract, sort, evolve, and analyze their importance, impact, and, thus finally manage knowledge in GSD projects. In particular, there is interest to consider the knowledge related to: domain (business rules), development process, technical resources (project management tool, version control system, among others), non-technical resources (human resources, social and cultural aspects) and technologies (programming language, database and others).
- To define mechanisms to support the capture of contextual information that can come from multiple and heterogeneous sources, since the scenario is GSD. It is also important to set the frequency with which captures must occur on each one of the data.
- To explore different cloud computing realizations (Platform as a Service, Infrastructure as a Service and Software as a Service). Particularly, our interest is the definition of adequate strategies to storage and management of data from different sites. The data are related to contextual information or metrics used for performance analysis, as well as from GSD projects. Knowledge will be extract from them.
- To establish appropriate ways for the representation, processing and dissemination of knowledge and contextual information. Among the various subdomains of interest in knowledge management are ontologies, which allow explain, formalize, and represent knowledge in a specific area. Some of the benefits of the use of ontology are: definition of consensual vocabulary; standardization of system development, possibility to developing smarter software associating inference techniques on ontology, and reuse of modeled knowledge in new applications and research.
- To study the techniques to enable the correct presentation, visualization and dissemination of contextual information according to the goal, profile and format.
- To design, adapt or develop an ontology for global software development domain, considering the representation of contextual information for multiple subdomains (technical, organizational and socio-cultural aspects) of the GSD.
- To map the socio-cultural aspects, captured in GSD projects, to each phase of knowledge conversion model (socialization, externalization, combination and internalization).
• To define the performance dimensions for decision-making process, and also for the continuous improvement of processes and products in GSD. Define the performance indicators and the performance management system from GSD. From the practice of performance measurement will be possible to design management mechanisms to support development, reuse and capitalization of knowledge.

• To provide the decision maker with a tool that offers the support needed to capture, recovery, reuse and integration of contextual information and knowledge generated by distributed teams.

• To establish closer relationship among academy and industry. With it, search by solutions with character of technological innovation can be developed, making possible the technology transfer from academy to industry.

• To strengthen the relationship among research groups in KM for GSD and so thereby maximizes opportunities for fundraising, as well as for consolidation of the area with multidisciplinary characteristics. It could help to find technological solutions that lead to the projection of the country as a whole.

V. FINAL CONSIDERATIONS

Several studies related to GSD, Knowledge Management and Contextual Information can be found published at important events and journals in Software Engineering area, such as: International Conference on Software Engineering(ICSE), Brazilian Symposium on Software Engineering (SBES), Brazilian Symposium on Software Quality (SBQS), International Congress on Global Software Engineering (ICGSE), IEEE Software, Communications of ACM, Information and Software Technology, ACM Queue, Journal of Brazilian Computer Science (JBCS), IET, among others. Large organizations have increasingly taken advantage of opportunities created by global competition in software development. Furthermore, the progress of the economy and the sophistication on communication media, coupled with pressures on costs and quality has encouraged investments in global software development. However, in the nature of GSD work becomes important to consider one more variable: the contextual information.

So the main contribution of this position paper is to emphasize the opportunities and research challenges in GSD, specially related to contextual information management, knowledge management and performance management. The research agenda can direct the academic and industrial interests, and also establish partnership among several research groups of area. This contributes to reinforce the area in Brazil with solutions for the challenges identified.

The experience acquired by our research group, the trends identified in the literature as well as the market behavior, emphasize that is important that efforts be made to provide new technologies and development approaches that take into account the complexity of software products. This increase in complexity and also in the number of software development organizations, generate a large volume of data that needs to be adequately dealt to result in knowledge. This knowledge can be used to add value to decision-making. Still, it can be observed also for organizations to become more competitive it is necessary to adopt practices and tools that provide support to reduce time and cost of software development.

Particularly in Brazil, it is possible to see an increase in the interest of research groups in GSD, which has led to an increased country visibility due to the cooperation established among Brazilian and foreign researchers.

Moreover, the last two decades knowledge came to be considered as one of most important assets to be managed and implemented by organizations to achieve competitive advantages against their concurrent. Therefore, it is possible to identify evidence of the need to develop approaches to deal with contextual information and knowledge management, especially in collaborative scenarios. These scenarios lead to the adoption of GSD approach and, thus, it is expected to achieve higher performance in software manufacturing.

ACKNOWLEDGMENT

This research has been sponsored by Araucária Fundation. The author Rafael Leonardo Vivian thanks CAPES for the financial support.

REFERENCES


